Generalized Residual Multiple Model Adaptive Estimation of Parameters and States

DISSERTATION

Charles D. Ormsby, Major, USAF

AFIT/DS/ENG/03-08

DEPARTMENT OF THE AIR FORCE

AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright Patterson Air Force Base, Ohio

Approved for public release; distribution unlimited

The views expressed in this dissertation are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government.

AFIT/DS/ENG/03-08

Generalized Residual Multiple Model Adaptive Estimation of Parameters and States

DISSERTATION

Presented to the Faculty of the Graduate School of Engineering and Management
of the Air Force Institute of Technology
Air University
In Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

Charles D. Ormsby, B.S., M.S.
Major, USAF

October 2, 2003

Approved for public release; distribution unlimited

AFIT/DS/ENG/03-08

Generalized Residual Multiple Model Adaptive Estimation
of Parameters and States

Charles D. Ormsby, B.S., M.S.

Major, USAF

Approved:

Dr. John F. Raquet Dissertation Advisor	9 OCT 03 Date
Dr. Peter S. Maybeck Committee Member	2 Oct 03 Date
Dr. Mark E. Oxley Committee Member	2 Oct 03 Date
Dr. Michael A. Temple Committee Member	2 oct 03 Date
Dr. Anthony N. Palazotto Dean's Representative	2 oct 03 Date

Robert A. Calico, Jr
Dean

Acknowledgments

I wish to thank my committee members, Dr. Oxley and Dr. Temple. Their outside points of view have been extremely valuable in helping me ensure I present my work in an understandable fashion. Additionally, they have repeatedly provided insights and suggestions that prevented me from getting "MMAE tunnel-vision." I thank them both for their dedication and hard work reviewing and critiquing this research. I also wish to thank Dr. Maybeck, my third committee member. He has been an inspiration throughout this research effort. His concern for me as a student and his desire to do what's best for my education have made him a joy to work with. His insistence on excellence and correctness, both in his work and mine, have made him a mentor that I admire deeply. People of Dr. Maybeck's stature are few and far between. I will continue to measure my work using the "Maybeck yardstick" for the rest of my career. Finally, the biggest thanks (at work) goes to my advisor Dr. Raquet. His immense intelligence astounds me. More importantly, his work ethic greatly overshadows his God-given ability. I'll never know exactly how many times he was up half the night reviewing my work so that he could return it to me quickly. I sometimes wonder if his work ethic drove him to put more effort into my education than I did. I am deeply indebted to him for his help along this educational journey, and I am inspired by the effort he puts in at AFIT everyday. I will repay him the only way I know how, by approaching my jobs with the same dedication that he put into teaching me. Thank you! Finally, I wish to thank the most important people in my life, my family. First, to my parents who taught me at a young age that I could accomplish anything with hard work. They recognized my academic abilities early on and convinced me that a higher education was the way to succeed in my career. Thanks to them for the confidence and for teaching me the importance of an education and hard work. Next, to my two little girls who reminded me daily that there's more to life than school. I'll never forget *very* early on a Saturday morning when my then 3 year old daughter came walking from her room to see what I was doing out of bed so early. I explained that I got up early to work on a report for school so that I would have time to play with her and her sister later in the day. I told her I was working on making some plots "pretty" for my report. In typical 3 year old fashion, she gleefully responded, "I know how to make them pretty for you Daddy." She then went

to her room, came back with some stickers, and we spent the next several minutes together making my plots "pretty." The results hung on my bulletin board at work for the next 2 years as a reminder of what truly matters in life. Thanks to both girls for the joy of nightly games of "jungle gym" and the pride of watching them grow. I love you both more than you'll ever know. Most importantly, I owe the deepest debt of gratitude to my wife. During our 8 years of marriage, you have always been unconditionally supportive of me. You have now experienced the life of the "AFIT widow" twice. When I suggested coming back to AFIT for the PhD program, despite knowing what it would mean to you, you agreed without hesitation. Anytime I needed to stay late or work over a weekend, your response was always the same, "Do whatever you need to do, we'll be fine." Now, as I finish up this degree, you're living a thousand miles away at our follow-on assignment so our daughter can start her educational journey off right. You are the rock of our family. You are a silent martyr, joyfully doing whatever is best for the other members of our family without regard for yourself. Honey, I can never express how deep my love for you runs. Thank you for all the sacrifices you make for me. I'm truly blessed to have found such a loving, supportive wife.

Notation

The following notation is used throughout this document. It is primarily influenced by [29] as well as having minor influence from other authors listed in the bibliography.

Vectors	lower case bold face letters such as $\mathbf{x}$
Matrices	upper case bold face letters such as $\mathbf{P}$
Random Vectors	lower case bold face san serif type such as x
Random Matrices	upper case bold face san serif type such as P
Realizations of Random Vectors and Matrices	lower and upper case, respectively, bold face roman type such as $\mathbf{x}$ and $\mathbf{P}$
Arguments of Densities, Integrals, etc.	Greek letter dummy variables such as ξ
Superscript T	transpose
Superscript -1	inverse
Superscript $+$	quantity after measurement update in the Kalman filter
Superscript $-$	quantity before measurement update in the Kalman filter
Raised "Hat" ($\widehat{\ }$)	estimate
$\lvert \cdot \rvert$	Matrix determinant
Superscript $*$	Quantity or variable associated with generalized residual (not the complex conjugate)

Table of Contents

Page

List of Figures

List of Tables

List of Symbols

Symbol Page

Symbol Page

List of Abbreviations

Abbreviation	Page
BPSK Binary Phase Shift Keying	2-30
FOGMA First Order Gauss-Markov Acceleration	4-1
GPS Global Positioning System	1-1
GRMMAE Generalized Residual Multiple Model Adaptive Estimator	1-4
IMM Interacting Multiple Model	2-29
INS Inertial Navigation System	4-17
JPALS Joint Precision Approach and Landing System	1-1
KDI Kullback Discrimination Index	2-27
MMAE Multiple Model Adaptive Estimator	1-2
PDF Probability Density Function	2-2
SA Selective Availability	2-31

Abstract

This dissertation develops a modification to the standard Multiple Model Adaptive Estimator (MMAE) which allows the use of a new "generalized residual" in the hypothesis conditional probability calculation. The generalized residual is a linear combination of traditional Kalman filter residuals and "post-fit" Kalman filter residuals which are calculated after measurement incorporation. This modified MMAE is termed a Generalized Residual Multiple Model Adaptive Estimator (GRMMAE). The dissertation provides a derivation of the hypothesis conditional probability formula which the GRMMAE uses to calculate probabilities that each elemental filter in the GRMMAE contains the correct parameter value. Through appropriate choice of a single scalar GRMMAE design parameter, the GRMMAE can be designed to be equivalent to a traditional MMAE, a post-fit residual modified MMAE, or any number of yet unused MMAEs. The original GRMMAE design goal was to choose the GRMMAE design parameter which caused the fastest GRMMAE convergence to the correct hypothesis. However, this dissertation demonstrates that the GRMMAE design parameter can lead to β-dominance, a negative performance effect in the GRMMAE. This is a key contribution of this research as other researchers have previously suggested that the use of post-fit residuals may be advantageous in certain MMAE applications. This dissertation directly addresses the use of post-fit residuals by those researchers and demonstrates that, for their application, equivalent performance is achieved using a traditional MMAE.

Generalized Residual Multiple Model Adaptive Estimation

of Parameters and States

I. Introduction

1.1 Overview

It is often necessary to estimate both a system parameter and the system state. A common example of this need is precise relative positioning using Global Positioning System (GPS) carrier-phase measurements. These precise GPS positions are routinely used in surveying and are planned for the US Navy variant of the Joint Precision Approach and Landing System (JPALS). The next chapter demonstrates that a constant parameter, namely carrier-phase integer ambiguities, must be estimated in order to obtain carrier-phase-based GPS positions. The parameter estimate is then used as part of the system model to estimate the system state variables (typically including position, velocity, and acceleration in a navigation system). Thus, there is a motivation to develop techniques for estimating states in the presence of an unknown system parameter. In general, these techniques use sensor measurements to estimate the system parameter. In turn, the parameter estimate is used to calculate the state variable estimates. A generic block diagram of such a system is shown in Figure 1.1 where $\mathbf{z}$ is the measurement vector, $\widehat{\mathbf{x}}$ is the state vector estimate, and $\widehat{\mathbf{a}}$ is the parameter vector estimate.

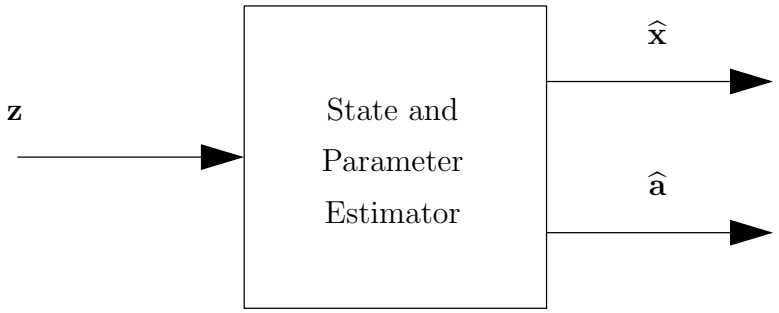


Figure 1.1 Generic Parameter and State Estimator

A second common application for the system in Figure 1.1 is failure detection and isolation. For this application, the system parameter is an operating mode (i.e., fully functional, type 1 failure, type 2 failure, etc.). The generic parameter and state estimator uses sensor measurements to estimate the system mode and to produce an appropriate state variable estimate.

By defining an appropriate system parameter, both of these common estimation applications can be addressed using the generic system of Figure 1.1. Since this research was inspired by GPS carrier-phase integer ambiguity resolution, that application is discussed in some detail. However, this research is sufficiently general to apply to the failure detection and isolation application as well as other appropriate applications where the system state is estimated in the presence of an unknown system parameter. Chapters II and III explicitly state all mathematical assumptions an application must meet for this research to apply.

The Multiple Model Adaptive Estimator (MMAE) [26] of Figure 1.2 is a specific type of parameter and state estimator. The MMAE consists of a bank of parallel Kalman filters, termed "elemental filters." Each elemental filter is based on a hypothesized parameter vector, $\mathbf{a}_1$, $\mathbf{a}_2$, ... or, $\mathbf{a}_K$, where $\mathbf{a}_j$ is the constant (or slowly-varying with respect to the states) parameter vector for the j^{th} elemental filter, and there are K elemental filters in the MMAE. The elemental filters act upon a measurement vector $\mathbf{z}$ to produce a state estimate $\widehat{\mathbf{x}}_j$, a residual vector $\mathbf{r}_j$, and a residual covariance matrix, $\mathbf{A}_j$. The MMAE is derived in detail in Chapter II.

The primary motivation for using a multiple filter technique like the MMAE rather than a single filter technique is responsiveness. In many cases, problems that are addressed using the MMAE could also be addressed with a single filter estimation technique such as a Kalman filter. However, by comparing measurement residuals, $\mathbf{r}$, from several elemental filters, the MMAE is generally able to determine the correct parameter value more quickly. Additionally, the MMAE will generally respond to changes in the parameter, such as a system failure, more quickly than a single filter solution [46]. The additional filters required in the MMAE do increase the computational load when compared to a single filter solution. However, computational power continues to increase, and computationally efficient

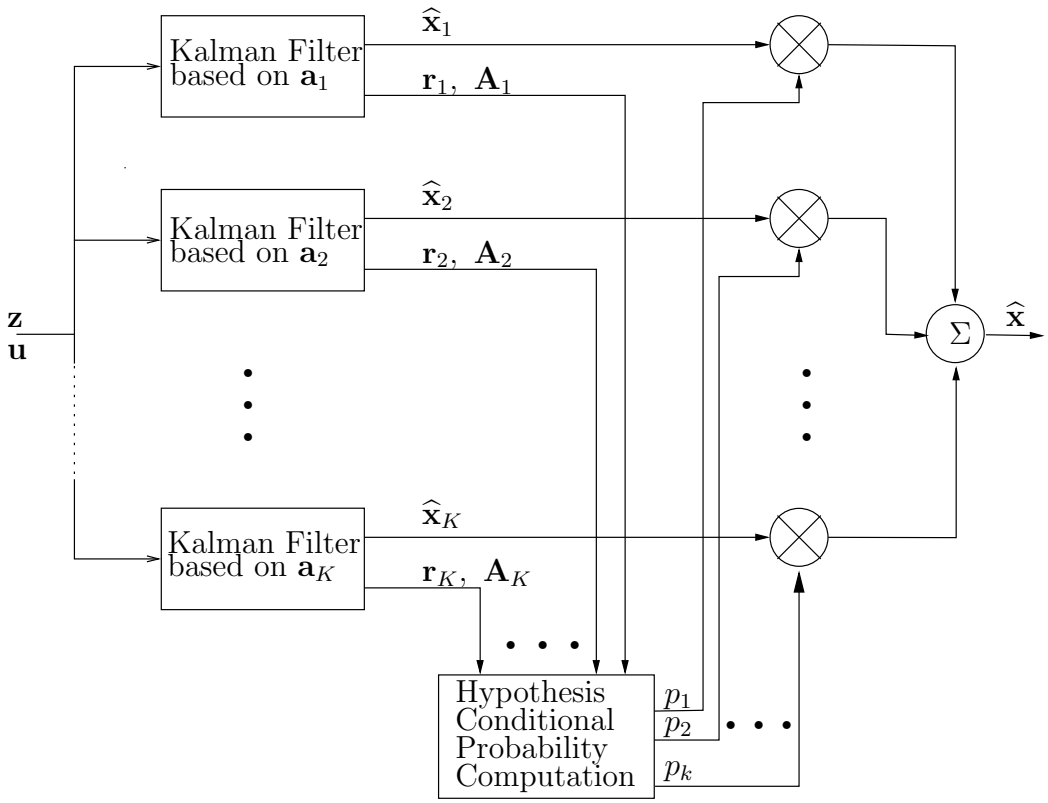


Figure 1.2 General MMAE Block Diagram

algorithms such as "moving-bank" and "filter spawning" MMAE's have been developed to address this disadvantage [10, 11, 12, 13, 14, 23, 24, 28, 41, 42, 46]. This research assumes that sufficient computational power exists, so methods to increase computational efficiency are not emphasized.

Recently, researchers have suggested modifying the MMAE to calculate hypothesis conditional probabilities from "post-fit" residuals [17, 18]. The traditional MMAE uses a nonlinear function of measurement residuals from each elemental filter to calculate probabilities that the filter contains the correct hypothesis. These researchers have suggested that better performance may be achieved by calculating hypothesis conditional probabilities using a nonlinear function of residuals calculated *after* measurement update. Such residuals are termed "post-fit" residuals. Additionally, [47] has suggested using post-fit residuals from multiple Kalman filters for hypothesis testing. However, it is unclear from [47] if a full MMAE was used or if post-fit residual monitoring was used. Regardless of

what technique [47] used, the method is at least closely related to using post-fit residuals in an MMAE.

This dissertation was inspired by [17, 18], which provided no formal derivation or mathematical characterization of the effect that post-fit residual processing has on the MMAE. A research contribution of this dissertation is to fill the gap left in [17, 18]. To do this, a mathematical generalization to the MMAE is developed to allow the designer easily to incorporate linear combinations of Kalman filter residuals (calculated before measurement update) and post-fit residuals (calculated after measurement update). This linear combination is termed a "generalized residual." The MMAE acting upon this generalized residual is introduced here as the "Generalized Residual Multiple Model Adaptive Estimator" (GRMMAE). This research provides a derivation of the hypothesis conditional probability calculation based upon the generalized residual. Additionally, generalized residual properties for both correct and incorrect hypotheses are derived. The effect of using the generalized residual on MMAE performance is demonstrated. The fact that the generalized residual leads to the β-dominance effect [2, 3, 32], a negative performance attribute, is demonstrated. Additionally, for the specific design and application used by [17, 18, 21], this dissertation demonstrates that the β-dominance effect is cancelled and proves that the suggestion in [17, 18] that post-fit residuals improve performance for GPS carrier-phase integer ambiguity resolution is not true. These are all contributions of this research. Further, this dissertation demonstrates that, in general, the proposed GRMMAE is susceptible to the β-dominance effect and is therefore, not recommended for wide application. However, as the recommendation for future research suggests, with further development, the GRMMAE may be useful to correct for β-dominance optimally.

1.2 Dissertation Summary

This dissertation consists of 5 chapters. This chapter provides an introduction to the dissertation. Chapter II gives necessary background for the new research. Chapter III derives the new theory for this dissertation, presenting a full derivation of the proposed GRMMAE and properties of the generalized residuals. The primary research contributions of this dissertation are contained in Chapter III. Chapter IV presents simulation and

analysis results for the GRMMAE, demonstrating that the modified MMAE from [17, 18] (a specific example of the GRMMAE) provides equivalent performance to the traditional MMAE. Additionally, it demonstrates that, in general, the proposed GRMMAE is susceptible to the β-dominance effect. Simulation results in Chapter IV validate theoretical contributions presented in Chapter III. Chapter V gives conclusions and recommendations based on this research. Additionally, the research contributions of this dissertation are enumerated in Chapter V.

II. Background

2.1 Kalman Filter

The basic building block in the Multiple Model Adaptive Estimator (MMAE) is the Kalman filter. Therefore, the Kalman filter is derived here. This section presents the linear Kalman filter used for systems which can be described by linear dynamics and measurement models. Section 2.2 describes the extended Kalman filter, which is used when a system has nonlinearities in the dynamics and/or measurement model.

Consider a system which can be modeled by the linear stochastic difference equation

$$\begin{aligned} \mathbf{x}(t_{i+1}) &= \mathbf{\Phi}(t_{i+1}, t_i)\,\mathbf{x}(t_i) + \mathbf{B}_d(t_i)\,\mathbf{u}(t_i) + \mathbf{w}_d(t_i) \\ \mathbf{z}(t_i) &= \mathbf{H}(t_i)\,\mathbf{x}(t_i) + \mathbf{v}(t_i) \end{aligned} \tag{2.1}$$

In the first line of Equation (2.1), $\mathbf{x}(t_i)$ is the state vector, $\mathbf{\Phi}(t_{i+1}, t_i)$ is the state transition matrix from time t_i to t_{i+1}, $\mathbf{B}_d(t_i)$ is the input matrix, $\mathbf{u}(t_i)$ is the control vector, and $\mathbf{w}_d(t_i)$ is the discrete-time dynamics driving noise vector. This represents the system dynamics model. The second line of Equation (2.1) gives the system measurement model, where $\mathbf{z}(t_i)$ is the measurement vector, $\mathbf{H}(t_i)$ is the measurement matrix, and $\mathbf{v}(t_i)$ is the measurement noise vector. Note that Equation (2.1) may be an equivalent discrete-time description of a continuous-time system model [30]. The form of Equation (2.1) assumes measurements $\mathbf{z}(t_i)$ are available only at discrete-times.

Noise vectors $\mathbf{w}_d(t_i)$ and $\mathbf{v}(t_i)$ are assumed independent, zero-mean, white, Gaussian processes with covariance kernels

$$\begin{aligned} E\left\{\mathbf{w}_d(t_i)\,\mathbf{w}_d^T(t_j)\right\} &= \mathbf{Q}_d(t_i)\,\delta_{ij} \\ E\left\{\mathbf{v}(t_i)\,\mathbf{v}^T(t_j)\right\} &= \mathbf{R}(t_i)\,\delta_{ij} \end{aligned} \tag{2.2}$$

where the discrete-time dynamics driving noise covariance $\mathbf{Q}_d(t_i)$ is assumed positive semi-definite for all time, the measurement noise covariance $\mathbf{R}(t_i)$ is assumed positive definite for

all time, and δ_{ij} is the Kronecker delta function. It is assumed the system of Equation (2.1) starts from an initial state $\mathbf{x}(t_0)$, which is a Gaussian random vector with mean and covariance given by

$$\begin{aligned} E\{\mathbf{x}(t_0)\} &= \widehat{\mathbf{x}}_0 \\ E\left\{[\mathbf{x}(t_0) - \widehat{\mathbf{x}}_0][\mathbf{x}(t_0) - \widehat{\mathbf{x}}_0]^T\right\} &= \mathbf{P}_0 \end{aligned} \tag{2.3}$$

where $\widehat{\mathbf{x}}_0$ is the initial state vector estimate, $\mathbf{x}(t_0)$ is assumed independent of $\mathbf{w}_d(t_i)$ and $\mathbf{v}(t_i)$ for all t_i, and $\mathbf{P}_0$ is the initial state covariance matrix. $\mathbf{P}_0$ is, by definition, positive semi-definite.

Two measurement times t_{i-1} and t_i are used to derive a two-step process. The first step propagates the optimal state estimate and state covariance from the time immediately after measurement update at t_{i-1} to the time immediately before measurement update at t_i. The next step updates the propagated state and covariance estimates using the measurement update at time t_i. Notationally, t_i^- is used to denote a quantity at time t_i immediately *before* the measurement update is incorporated, while t_i^+ denotes a quantity immediately *after* the measurement update at time t_i. The derived propagate and update algorithm for the general times t_{i-1} and t_i is used recursively to estimate the state vector and state covariance, beginning at the initial conditions of Equation (2.3).

Suppose the measurement $\mathbf{z}(t_{i-1})$ has been incorporated. The desire is to define the conditional probability density function (PDF) for the state vector conditioned on the entire measurement history $f_{\mathbf{x}(t_{i-1})|\mathbf{Z}(t_{i-1})}(\boldsymbol{\xi}|\mathbf{Z}(t_{i-1}) = \boldsymbol{\mathcal{Z}}_{i-1})$, and to propagate that density forward until the next measurement is incorporated. Here, $\mathbf{Z}(t_{i-1})$ is a growing length measurement history vector consisting of all measurement vectors from t_0 through t_{i-1}, $\boldsymbol{\xi}$ and $\boldsymbol{\mathcal{Z}}_{i-1}$ are dummy variables representing the state vector and the measurement history, respectively. It is proven in [29] that $f_{\mathbf{x}(t_{i-1})|\mathbf{Z}(t_{i-1})}(\boldsymbol{\xi}|\mathbf{Z}(t_{i-1}) = \boldsymbol{\mathcal{Z}}_{i-1})$ is Gaussian under the previous assumptions concerning Equations (2.1) through (2.3). Therefore, the mean and covariance fully define the density. For convenience, define

$$\widehat{\mathbf{x}}\left(t_{i-1}^{+}\right) \triangleq E\Big[\mathbf{x}\left(t_{i-1}\right) \mid \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\Big]$$
$$\mathbf{P}\left(t_{i-1}^{+}\right) \triangleq E\Big[\left[\mathbf{x}\left(t_{i-1}\right) - \widehat{\mathbf{x}}\left(t_{i-1}^{+}\right)\right]\left[\mathbf{x}\left(t_{i-1}\right) - \widehat{\mathbf{x}}\left(t_{i-1}^{+}\right)\right]^{T} \mid \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\Big] \quad (2.4)$$

where $\widehat{\mathbf{x}}\left(t_{i-1}^{+}\right)$ and $\mathbf{P}\left(t_{i-1}^{+}\right)$ are the conditional state mean and covariance estimates, conditioned on the realized measurement history through time t_{i-1}, $\mathbf{Z}_{i-1}$, immediately after measurement incorporation at t_{i-1}. For a Gaussian density, the conditional mean $\widehat{\mathbf{x}}\left(t_{i-1}^{+}\right)$ is the optimal state estimate following measurement update at time t_{i-1} for "essentially any criterion of optimality" [29]. Not only is it the conditional mean, it is also the conditional mode and the conditional median.

Next, the conditional mean and covariance need to be propagated to time t_i^-. To simplify this discussion, assume there are no control inputs (i.e., $\mathbf{u}\left(t_{i-1}\right) = \mathbf{0}$). Control inputs only have the effect of shifting the state conditional mean. They have no effect on the conditional covariance. The effect of control inputs is added following this development. The state conditional mean at t_i, conditioned on the measurement history at t_{i-1} is

$$\begin{aligned} E\Big[\mathbf{x}\left(t_i\right) \mid \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\Big] &= E\Big[\mathbf{\Phi}\left(t_i, t_{i-1}\right)\mathbf{x}\left(t_{i-1}\right) + \mathbf{w}_d\left(t_{i-1}\right) \mid \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\Big] \\ &= \mathbf{\Phi}\left(t_i, t_{i-1}\right) E\Big[\mathbf{x}\left(t_{i-1}\right) \mid \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\Big] \\ &\quad + E\Big[\mathbf{w}_d\left(t_{i-1}\right)\Big] \\ E\Big[\mathbf{x}\left(t_i\right) \mid \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\Big] &= \mathbf{\Phi}\left(t_i, t_{i-1}\right) E\Big[\mathbf{x}\left(t_{i-1}\right) \mid \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\Big] \end{aligned} \quad (2.5)$$

In the third line of Equation (2.5), the unconditional mean is used because $\mathbf{w}_d\left(t_i\right)$ is independent of $\mathbf{Z}\left(t_{i-1}\right)$. In the final line, the assumption that $\mathbf{w}_d\left(t_i\right)$ is zero-mean is invoked.

Define

$$\begin{aligned}
\widehat{\mathbf{x}}\left(t_i^-\right) &\triangleq E\Big[\mathbf{x}\left(t_i\right) \mid \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\Big] \\
\mathbf{P}\left(t_i^-\right) &\triangleq E\Big[\left[\mathbf{x}\left(t_i\right) - \widehat{\mathbf{x}}\left(t_i^-\right)\right]\left[\mathbf{x}\left(t_i\right) - \widehat{\mathbf{x}}\left(t_i^-\right)\right]^T \mid \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\Big]
\end{aligned} \tag{2.6}$$

Using Equations (2.5) and (2.6), the propagated state conditional mean is

$$\widehat{\mathbf{x}}\left(t_i^-\right) = \mathbf{\Phi}\left(t_i, t_{i-1}\right)\widehat{\mathbf{x}}\left(t_{i-1}^+\right) \tag{2.7}$$

The propagated state covariance estimate is

$$\begin{aligned}
\mathbf{P}\left(t_i^-\right) &= E\Big[\left[\mathbf{x}\left(t_i\right) - \widehat{\mathbf{x}}\left(t_i^-\right)\right]\left[\mathbf{x}\left(t_i\right) - \widehat{\mathbf{x}}\left(t_i^-\right)\right]^T \mid \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\Big] \\
&= E\Big[\left[\mathbf{\Phi}\left(t_i, t_{i-1}\right)\mathbf{x}\left(t_{i-1}\right) + \mathbf{w}_d\left(t_i\right) - \widehat{\mathbf{x}}\left(t_i^-\right)\right]\left[\mathbf{\Phi}\left(t_i, t_{i-1}\right)\mathbf{x}\left(t_{i-1}\right) + \mathbf{w}_d\left(t_i\right)\right. \\
&\quad \left. - \widehat{\mathbf{x}}\left(t_i^-\right)\right]^T \mid \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\Big] \\
&= \mathbf{\Phi}\left(t_i, t_{i-1}\right) E\Big[\mathbf{x}\left(t_{i-1}\right)\mathbf{x}^T\left(t_{i-1}\right) \mid \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\Big]\mathbf{\Phi}^T\left(t_i, t_{i-1}\right) \\
&\quad - \widehat{\mathbf{x}}\left(t_i^-\right) E\Big[\mathbf{x}^T\left(t_{i-1}\right) \mid \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\Big]\mathbf{\Phi}^T\left(t_i, t_{i-1}\right) \\
&\quad - \mathbf{\Phi}\left(t_i, t_{i-1}\right) E\Big[\mathbf{x}\left(t_{i-1}\right) \mid \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\Big]\widehat{\mathbf{x}}^T\left(t_i^-\right) \\
&\quad + \widehat{\mathbf{x}}\left(t_i^-\right)\widehat{\mathbf{x}}^T\left(t_i^-\right) + E\Big[\mathbf{w}_d\left(t_i\right)\mathbf{w}_d^T\left(t_i\right) \mid \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\Big] \\
&= \mathbf{\Phi}\left(t_i, t_{i-1}\right) E\Big[\mathbf{x}\left(t_{i-1}\right)\mathbf{x}^T\left(t_{i-1}\right) \mid \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\Big]\mathbf{\Phi}^T\left(t_i, t_{i-1}\right) \\
&\quad - \widehat{\mathbf{x}}\left(t_i^-\right)\widehat{\mathbf{x}}^T\left(t_{i-1}^+\right)\mathbf{\Phi}^T\left(t_i, t_{i-1}\right) - \mathbf{\Phi}\left(t_i, t_{i-1}\right)\widehat{\mathbf{x}}\left(t_{i-1}^+\right)\widehat{\mathbf{x}}^T\left(t_i^-\right) + \widehat{\mathbf{x}}\left(t_i^-\right)\widehat{\mathbf{x}}^T\left(t_i^-\right) \\
&\quad + \mathbf{Q}_d\left(t_i\right) \\
&= \mathbf{\Phi}\left(t_i, t_{i-1}\right) E\Big[\mathbf{x}\left(t_{i-1}\right)\mathbf{x}^T\left(t_{i-1}\right) \mid \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\Big]\mathbf{\Phi}^T\left(t_i, t_{i-1}\right) \\
&\quad - \mathbf{\Phi}\left(t_i, t_{i-1}\right)\widehat{\mathbf{x}}\left(t_{i-1}^+\right)\widehat{\mathbf{x}}^T\left(t_{i-1}^+\right)\mathbf{\Phi}^T\left(t_i, t_{i-1}\right) \\
&\quad - \mathbf{\Phi}\left(t_i, t_{i-1}\right)\widehat{\mathbf{x}}\left(t_{i-1}^+\right)\widehat{\mathbf{x}}^T\left(t_{i-1}^+\right)\mathbf{\Phi}^T\left(t_i, t_{i-1}\right) \\
&\quad + \mathbf{\Phi}\left(t_i, t_{i-1}\right)\widehat{\mathbf{x}}\left(t_{i-1}^+\right)\widehat{\mathbf{x}}^T\left(t_{i-1}^+\right)\mathbf{\Phi}^T\left(t_i, t_{i-1}\right) + \mathbf{Q}_d\left(t_i\right)
\end{aligned}$$

$$\begin{aligned}
\mathbf{P}\left(t_i^-\right) &= \mathbf{\Phi}\left(t_i, t_{i-1}\right)\left\{E\left[\mathbf{x}\left(t_{i-1}\right)\mathbf{x}^T\left(t_{i-1}\right)|\mathbf{Z}\left(t_{i-1}\right)=\mathbf{Z}_{i-1}\right]-\widehat{\mathbf{x}}\left(t_{i-1}^+\right)\widehat{\mathbf{x}}^T\left(t_{i-1}^+\right)\right\}\mathbf{\Phi}^T\left(t_i, t_{i-1}\right) \\
&\quad + \mathbf{Q}_d\left(t_i\right) \\
\mathbf{P}\left(t_i^-\right) &= \mathbf{\Phi}\left(t_i, t_{i-1}\right)\mathbf{P}\left(t_{i-1}^+\right)\mathbf{\Phi}\left(t_i, t_{i-1}\right) + \mathbf{Q}_d\left(t_i\right) \qquad (2.8)
\end{aligned}$$

where the cross-terms involving $\mathbf{x}\left(t_i\right)$ with $\mathbf{w}_d\left(t_i\right)$ were dropped because $\mathbf{w}_d\left(t_i\right)$ is assumed to be independent of $\mathbf{x}\left(t_i\right)$ and zero-mean. Similarly, the cross-terms with $\widehat{\mathbf{x}}\left(t_{i-1}^+\right)$ and $\mathbf{w}_d\left(t_i\right)$ were dropped because $\mathbf{w}_d\left(t_i\right)$ is zero-mean.

If $\widehat{\mathbf{x}}\left(t_i^-\right)$ is used as the optimal propagated state estimate prior to measurement update at t_i, then the error associated with this estimate is $\left[\mathbf{x}\left(t_i\right) - \widehat{\mathbf{x}}\left(t_i^-\right)\right]$. Since this error is zero-mean [29], $\mathbf{P}\left(t_i^-\right)$ is the estimation error covariance as well as the propagated state estimate covariance. Note also that the measurement vector $\mathbf{z}\left(t_i\right)$ does not appear in Equation (2.8). Therefore, given a covariance for the initial state estimate, the propagated state estimate covariance (and state estimate error covariance) can be computed for all time before any measurements are taken.

Now, consider updating the propagated state estimate with information from the measurement taken at t_i. Thus, the conditional density $f_{\mathbf{x}(t_i)|\mathbf{Z}(t_i)}\left(\boldsymbol{\xi}|\mathbf{Z}\left(t_i\right) = \boldsymbol{\mathcal{Z}}_i\right)$ is desired. Repeated application of Baye's rule and the definition of the measurement history yield [29]

$$\begin{aligned}
&f_{\mathbf{x}(t_i)|\mathbf{Z}(t_i)}\left(\boldsymbol{\xi}|\mathbf{Z}\left(t_i\right) = \boldsymbol{\mathcal{Z}}_i\right) \\
&\quad = \frac{f_{\mathbf{z}(t_i)|\mathbf{x}(t_i),\mathbf{Z}(t_{i-1})}\left(\boldsymbol{\zeta}_i|\mathbf{x}\left(t_i\right)=\boldsymbol{\xi}, \mathbf{Z}\left(t_{i-1}\right)=\boldsymbol{\mathcal{Z}}_{i-1}\right) f_{\mathbf{x}(t_i)|\mathbf{Z}(t_{i-1})}\left(\boldsymbol{\xi}|\mathbf{Z}\left(t_{i-1}\right) = \boldsymbol{\mathcal{Z}}_{i-1}\right)}{f_{\mathbf{z}(t_i)|\mathbf{Z}(t_{i-1})}\left(\boldsymbol{\zeta}_i|\mathbf{Z}\left(t_{i-1}\right) = \boldsymbol{\mathcal{Z}}_{i-1}\right)}
\end{aligned} \qquad (2.9)$$

where $\boldsymbol{\zeta}_i$ is the dummy variable associated with $\mathbf{z}\left(t_i\right)$.

The form of Equation (2.9) is motivated by the fact that all three densities on the right hand side are easily evaluated. The density $f_{\mathbf{x}(t_i)|\mathbf{Z}(t_{i-1})}\left(\boldsymbol{\xi}|\mathbf{Z}\left(t_{i-1}\right) = \boldsymbol{\mathcal{Z}}_{i-1}\right)$ is the density for the propagated state, developed earlier in this section. It is Gaussian [29], with mean and covariance given by Equations (2.7) and (2.8):

$$f_{\mathbf{x}(t_i)|\mathbf{Z}(t_{i-1})}\left(\boldsymbol{\xi}|\mathbf{Z}\left(t_{i-1}\right) \;=\; \boldsymbol{\mathcal{Z}}_{i-1}\right)$$
$$= \frac{1}{(2\pi)^{n/2}\left|\mathbf{P}\left(t_i^-\right)\right|^{1/2}} \exp\left\{-\frac{1}{2}\left[\boldsymbol{\xi}-\widehat{\mathbf{x}}\left(t_i^-\right)\right]^T \mathbf{P}^{-1}\left(t_i^-\right)\left[\boldsymbol{\xi}-\widehat{\mathbf{x}}\left(t_i^-\right)\right]\right\} \quad (2.10)$$

where n is the dimension of the state vector and $|\cdot|$ is the matrix determinant.

Next, consider the first conditional density in the numerator of Equation (2.9), $f_{\mathbf{z}(t_i)|\mathbf{x}(t_i),\mathbf{Z}(t_{i-1})}\left(\boldsymbol{\zeta}_i|\mathbf{x}\left(t_i\right)=\boldsymbol{\xi},\mathbf{Z}\left(t_{i-1}\right)=\boldsymbol{\mathcal{Z}}_{i-1}\right)$. Recall from Equation (2.1) that the measurement model is

$$\mathbf{z}\left(t_i\right) \;=\; \mathbf{H}\left(t_i\right)\mathbf{x}\left(t_i\right) \;+\; \mathbf{v}\left(t_i\right)$$

Since this density is conditioned on the fact the state has taken a value, $\mathbf{x}\left(t_i\right) \;=\; \boldsymbol{\xi}$, then the random vector $\mathbf{z}\left(t_i\right)$ is a linear function of the known state realization and the zero-mean, white, Gaussian measurement noise with covariance $\mathbf{R}\left(t_i\right)$. Therefore, this conditional density is a Gaussian density, with mean $\mathbf{H}\left(t_i\right)\boldsymbol{\xi}$ and covariance $\mathbf{R}\left(t_i\right)$:

$$f_{\mathbf{z}(t_i)|\mathbf{x}(t_i),\mathbf{Z}(t_{i-1})}\left(\boldsymbol{\zeta}_i|\mathbf{x}\left(t_i\right) \;=\; \xi,\; \mathbf{Z}\left(t_{i-1}\right) \;=\; \boldsymbol{\mathcal{Z}}_{i-1}\right)$$
$$= \frac{1}{(2\pi)^{m/2}\left|\mathbf{R}\left(t_i\right)\right|^{1/2}} \exp\left\{-\frac{1}{2}\left[\boldsymbol{\zeta}_i \;-\; \mathbf{H}\left(t_i\right)\boldsymbol{\xi}\right]^T \mathbf{R}^{-1}\left(t_i\right)\left[\boldsymbol{\zeta}_i \;-\; \mathbf{H}\left(t_i\right)\boldsymbol{\xi}\right]\right\} \quad (2.11)$$

where m is the dimension of the measurement vector.

Finally, consider $f_{\mathbf{z}(t_i)|\mathbf{Z}(t_{i-1})}\left(\boldsymbol{\zeta}_i|\mathbf{Z}\left(t_{i-1}\right) \;=\; \boldsymbol{\mathcal{Z}}_{i-1}\right)$, from Equation (2.9), which is Gaussian with mean $\mathbf{H}\left(t_i\right)\widehat{\mathbf{x}}\left(t_i^-\right)$ and covariance $\mathbf{A}\left(t_i^-\right)=\mathbf{H}\left(t_i\right)\mathbf{P}\left(t_i^-\right)\mathbf{H}^T\left(t_i\right)+\mathbf{R}\left(t_i\right)$, calculated from the measurement model in Equation (2.1) [29]:

$$f_{\mathbf{z}(t_i)|\mathbf{Z}(t_{i-1})}\left(\boldsymbol{\zeta}_i|\mathbf{Z}\left(t_{i-1}\right) \;=\; \boldsymbol{\mathcal{Z}}_{i-1}\right)$$
$$= \frac{1}{(2\pi)^{m/2}\left|\mathbf{A}\left(t_i^-\right)\right|^{1/2}} \exp\left\{-\frac{1}{2}\left[\boldsymbol{\zeta}_i \;-\; \mathbf{H}\left(t_i\right)\widehat{\mathbf{x}}\left(t_i^-\right)\right]^T \mathbf{A}^{-1}\left(t_i^-\right)\left[\boldsymbol{\zeta}_i \;-\; \mathbf{H}\left(t_i\right)\widehat{\mathbf{x}}\left(t_i^-\right)\right]\right\} \quad (2.12)$$

Substituting Equations (2.10) through (2.12) into Equation (2.9), and invoking the matrix inversion lemma, yields the equivalent Gaussian density [29]

$$\begin{aligned} & f_{\mathbf{x}(t_i)|\mathbf{Z}(t_i)}\left(\boldsymbol{\xi}|\mathbf{Z}\left(t_i\right) \;=\; \boldsymbol{\mathcal{Z}}_i\right) \\ & \quad = \;\frac{1}{\left(2\pi\right)^{n/2}\left|\mathbf{P}\left(t_i^+\right)\right|^{1/2}}\exp\{-\frac{1}{2}\left[\boldsymbol{\xi}-\widehat{\mathbf{x}}\left(t_i^+\right)\right]^T\mathbf{P}^{-1}\left(t_i^+\right)\left[\boldsymbol{\xi}-\widehat{\mathbf{x}}\left(t_i^+\right)\right]\} \end{aligned} \tag{2.13}$$

where the updated state vector mean and covariance immediately after measurement incorporation are

$$\begin{aligned} \widehat{\mathbf{x}}\left(t_i^+\right) \;&= \widehat{\mathbf{x}}\left(t_i^-\right) \;+\; \mathbf{P}\left(t_i^-\right)\mathbf{H}^T\left(t_i\right)\left[\mathbf{H}\left(t_i\right)\mathbf{P}\left(t_i^-\right)\mathbf{H}^T\left(t_i\right) \;+\; \mathbf{R}\left(t_i\right)\right]^{-1}\left[\mathbf{z}_i \;-\; \mathbf{H}\left(t_i\right)\widehat{\mathbf{x}}\left(t_i^-\right)\right] \\ \mathbf{P}\left(t_i^+\right) \;&= \mathbf{P}\left(t_i^-\right) \;-\; \mathbf{P}\left(t_i^-\right)\mathbf{H}^T\left(t_i\right)\left[\mathbf{H}\left(t_i\right)\mathbf{P}\left(t_i^-\right)\mathbf{H}^T\left(t_i\right) \;+\; \mathbf{R}\left(t_i\right)\right]^{-1}\mathbf{H}\left(t_i\right)\mathbf{P}\left(t_i^-\right) \end{aligned} \tag{2.14}$$

and $\mathbf{z}_i$ is the realized measurement at time t_i.

The Kalman filter propagate and update algorithm has now been developed assuming no control inputs are present. To complete the development, the effect of control inputs must be included. As mentioned previously, these deterministic inputs only affect the mean of the propagated state estimate. When control inputs are included, the propagated state vector estimate becomes [29]

$$\widehat{\mathbf{x}}\left(t_i^-\right) \;= \boldsymbol{\Phi}\left(t_i, t_{i-1}\right)\widehat{\mathbf{x}}\left(t_{i-1}^+\right) \;+\; \mathbf{B}_d\left(t_i\right)\mathbf{u}\left(t_i\right) \tag{2.15}$$

where the control input terms, from Equation (2.1), are the equivalent discrete-time control model calculated from the continuous-time model using

$$\mathbf{B}_d\left(t_i\right)\mathbf{u}\left(t_i\right) \;= \int_{t_{i-1}}^{t_i}\boldsymbol{\Phi}\left(t_i,\tau\right)\mathbf{B}\left(\tau\right)\mathbf{u}\left(\tau\right)d\tau \tag{2.16}$$

The Kalman filter is summarized by the propagation and update algorithm shown in Table 2.1. In Table 2.1, $\mathbf{K}(t_i)$ is the Kalman filter gain. All other variables and equations are previously defined in this section.

2.2 Extended Kalman Filter

Frequently, applications call for use of a Kalman filter on a nonlinear system. Such a system may contain nonlinearities in the dynamics and/or the measurement model. One example with a nonlinear measurement model is estimation of position using the GPS. A common solution for these nonlinear problems is the extended Kalman filter. Like the Kalman filter of Section 2.1, the extended Kalman filter can be a basic building block in the MMAE. Therefore, the extended Kalman filter is derived here, paralleling the derivation in [30]. It is important to note, however, that the MMAE based on extended Kalman filters is an ad hoc form. Unlike the MMAE based on linear system models and linear Kalman filters, it is not an optimal Bayesian form.

Consider a system which can be modeled by the continuous-time dynamics and discrete-time measurement models

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{f}[\mathbf{x}(t),\ \mathbf{u}(t),\ t] \ +\ \mathbf{G}(t)\,\mathbf{w}(t) \\ \mathbf{z}(t_i) &= \mathbf{h}[\mathbf{x}(t_i), t_i] \ +\ \mathbf{v}(t_i) \end{aligned} \tag{2.17}$$

Table 2.1 Kalman Filter Algorithm Summary [29]

State Propagation

$$\begin{aligned} \widehat{\mathbf{x}}\left(t_i^-\right) &= \mathbf{\Phi}(t_i, t_{i-1})\,\widehat{\mathbf{x}}\left(t_{i-1}^+\right) \ +\ \mathbf{B}_d(t_i)\,\mathbf{u}(t_i) \\ \mathbf{P}\left(t_i^-\right) &= \mathbf{\Phi}(t_i, t_{i-1})\,\mathbf{P}\left(t_{i-1}^+\right)\mathbf{\Phi}(t_i, t_{i-1}) \ +\ \mathbf{Q}_d(t_i) \end{aligned}$$

Measurement Update

$$\begin{aligned} \mathbf{K}(t_i) &= \mathbf{P}\left(t_i^-\right)\mathbf{H}^T(t_i)\left[\mathbf{H}(t_i)\,\mathbf{P}\left(t_i^-\right)\mathbf{H}^T(t_i) \ +\ \mathbf{R}(t_i)\right]^{-1} \\ \widehat{\mathbf{x}}\left(t_i^+\right) &= \widehat{\mathbf{x}}\left(t_i^-\right) \ +\ \mathbf{K}(t_i)\left[\mathbf{z}_i \ -\ \mathbf{H}(t_i)\,\widehat{\mathbf{x}}\left(t_i^-\right)\right] \\ \mathbf{P}\left(t_i^+\right) &= \mathbf{P}\left(t_i^-\right) \ -\ \mathbf{K}(t_i)\,\mathbf{H}(t_i)\,\mathbf{P}\left(t_i^-\right) \end{aligned}$$

where

$\dot{\mathbf{x}}(t)$	is the state vector derivative at time t
$\mathbf{f}[\cdot,\cdot,\cdot]$	is the general, possibly nonlinear, dynamics model; a function of $\mathbf{x}$, $\mathbf{u}$ and t
$\mathbf{h}[\cdot,\cdot]$	is the general, possibly nonlinear, measurement model; a function of $\mathbf{x}$ and t_i

For this derivation, it is assumed that $\mathbf{f}[\mathbf{x}(t),\ \mathbf{u}(t),\ t]$ is Lipschitz in $\mathbf{x}(t)$, continuous in $\mathbf{u}(t)$, and piecewise continuous in t. Additionally, the control input $\mathbf{u}(t)$ is assumed to be piecewise continuous in t.[1]

Assume $\mathbf{w}(t)$ and $\mathbf{v}(t_i)$ are independent, zero-mean, white, Gaussian noise processes with covariance kernels

$$\begin{aligned} E\left\{\mathbf{w}(t)\,\mathbf{w}^T(t\ +\ \tau)\right\} &= \mathbf{Q}(t)\,\delta(\tau) \\ E\left\{\mathbf{v}(t_i)\,\mathbf{v}^T(t_j)\right\} &= \mathbf{R}(t_i)\,\delta_{ij} \end{aligned} \tag{2.18}$$

where $\mathbf{Q}(t)$ is the continuous-time dynamics driving noise strength and $\delta(\tau)$ is the Dirac delta function. Finally, the model is assumed to start from initial state $\mathbf{x}(t_0)$, which is a Gaussian random vector with mean $\widehat{\mathbf{x}}_0$ and covariance $\mathbf{P}_0$. Note that in the case of a nonlinear dynamics model $\mathbf{f}[\mathbf{x}(t),\ \mathbf{u}(t),\ t]$, the state vector generally does not remain Gaussian beyond the initial time.

As a stepping stone to deriving the extended Kalman filter, first consider the linearized Kalman filter. The linearized Kalman filter assumes a nominal state trajectory $\mathbf{x}_n(t)$ is defined for all times of interest (defined as $t \in T$). Linear perturbation theory is then used to calculate state trajectory perturbations from the nominal trajectory. This theory is derived first, then the extended Kalman filter is related to it.

Assume the nominal state trajectory, subject to the initial condition $\mathbf{x}_n(t_0) = \mathbf{x}_{n_0}$, can be generated such that it satisfies the deterministic differential equation:

[1] These assumptions are necessary to ensure that a solution exists to the deterministic differential equation $\dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t),\ \mathbf{u}(t),\ t]$ subject to the initial condition $\mathbf{x}(t_0) = \mathbf{x}_0$. For more details, see [29] Section 2.3.

$$\dot{\mathbf{x}}_n(t) \;=\; \mathbf{f}\left[\mathbf{x}_n(t),\ \mathbf{u}(t), t\right] \tag{2.19}$$

To maintain the adequacy of linear perturbation techniques, perturbations of the control vector $\mathbf{u}(t)$ from a nominal control vector $\mathbf{u}_n(t)$ are not considered here.

This nominal state process has a set of nominal noise-free measurements $\mathbf{z}_n(t_i)$ associated with it, given by

$$\mathbf{z}_n(t_i) \;=\; \mathbf{h}\left[\mathbf{x}_n(t_i), t_i\right] \tag{2.20}$$

In Equations (2.19) and (2.20), $\mathbf{f}[\cdot,\cdot,\cdot]$ and $\mathbf{h}[\cdot,\cdot]$ are the general dynamics and measurement models, respectively, given in Equation (2.17).

Consider the state perturbation from the nominal state. Using Equations (2.17) and (2.19), the state perturbation from the nominal state is given by

$$\left[\dot{\mathbf{x}}(t) \;-\; \dot{\mathbf{x}}_n(t)\right] \;=\; \mathbf{f}\left[\mathbf{x}(t),\ \mathbf{u}(t), t\right] \;-\; \mathbf{f}\left[\mathbf{x}_n(t),\ \mathbf{u}(t), t\right] \;+\; \mathbf{G}(t)\,\mathbf{w}(t) \tag{2.21}$$

Note in Equation (2.21) that $\dot{\mathbf{x}}(t)$ and $\mathbf{x}(t)$ are random vectors while $\dot{\mathbf{x}}_n(t)$ and $\mathbf{x}_n(t)$ are deterministic.

To develop the perturbation equation, the right hand side term in Equation (2.21) is expanded using a Taylor series expansion about the nominal state vector $\mathbf{x}_n(t)$, yielding

$$\mathbf{f}\left[\mathbf{x}(t),\ \mathbf{u}(t), t\right] \;=\; \mathbf{f}\left[\mathbf{x}_n(t),\ \mathbf{u}(t), t\right] \;+\; \left.\frac{\partial \mathbf{f}\left[\mathbf{x},\ \mathbf{u}(t), t\right]}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_n} \left[\mathbf{x}(t) \;-\; \mathbf{x}_n(t)\right] \;+\; \text{h.o.t.} \tag{2.22}$$

In Equation (2.22), h.o.t. represents higher order terms and the derivative is assumed to exist. It is assumed that the state perturbations from nominal are sufficiently small to allow h.o.t. to be neglected. Substituting Equation (2.22) into Equation (2.21) yields a first-order linear approximation to the state perturbation from nominal

$$\dot{\boldsymbol{\delta}\mathbf{x}}(t) \;=\; \mathbf{F}\left[t;\mathbf{x}_n(t)\right]\boldsymbol{\delta}\mathbf{x}(t) \;+\; \mathbf{G}(t)\,\mathbf{w}(t) \tag{2.23}$$

where $\boldsymbol{\delta}\mathbf{x}(t)$ is a first-order approximation of $[\mathbf{x}(t) \;-\; \mathbf{x}_n(t)]$ and $\mathbf{F}\left[t;\mathbf{x}_n(t)\right]$ is a matrix of partial derivatives of the dynamics model defined by

$$\mathbf{F}\left[t;\mathbf{x}_n(t)\right] \;\triangleq\; \left.\frac{\partial \mathbf{f}\left[\mathbf{x},\;\mathbf{u}(t),t\right]}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_n} \tag{2.24}$$

Thus, a linear approximation for the perturbation of state vector from the nominal state vector is defined. To calculate the total state vector estimate, the perturbation estimate (to be derived) is added to the nominal state: $\widehat{\mathbf{x}}(t) \;=\; \widehat{\boldsymbol{\delta}\mathbf{x}}(t) \;+\; \mathbf{x}_n(t)$.

The same approach used to derive Equation (2.23) is used to derive the measurement perturbation

$$\boldsymbol{\delta}\mathbf{z}(t_i) \;=\; \mathbf{H}\left[t_i;\mathbf{x}_n(t_i)\right]\boldsymbol{\delta}\mathbf{x}(t_i) \;+\; \mathbf{v}(t_i) \tag{2.25}$$

As before, $\boldsymbol{\delta}\mathbf{z}(t_i)$ is a first-order approximation to $[\mathbf{z}(t_i) \;-\; \mathbf{z}_n(t_i)]$ and $\mathbf{H}\left[t_i;\mathbf{x}_n(t_i)\right]$ is a matrix of partial derivatives (assumed to exist) of the measurement model defined by

$$\mathbf{H}\left[t_i;\mathbf{x}_n(t_i)\right] \;\triangleq\; \left.\frac{\partial \mathbf{h}\left[\mathbf{x},t_i\right]}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_n(t_i)} \tag{2.26}$$

Equations (2.23) and (2.25) yield a linear model (albeit an approximation) to which linear filtering theory can be applied. This form of Kalman filter is called a *linearized Kalman filter* or *perturbation Kalman filter*. The extended Kalman filter is directly related to the linearized Kalman filter. The difference between the two lies in the fact that the extended Kalman filter re-linearizes about the state estimate after each measurement update, rather than linearizing about a pre-defined nominal state trajectory. The extended Kalman filter has the advantage over the linearized Kalman filter of being more robust in

the presence of large perturbations from the nominal state trajectory. For the extended Kalman filter to operate properly, though, any perturbations between two consecutive measurement updates must be small enough to allow h.o.t. from Equation (2.22) to be neglected.

To derive the extended Kalman filter, assume a measurement update has taken place at time t_i, producing state estimate $\widehat{\mathbf{x}}\left(t_i^+\right)$. Now, rather than maintaining the linearization about the nominal state $\mathbf{x}_n\left(t_i\right)$, re-linearize about the new state estimate. Let $\mathbf{x}_n\left(t/t_i\right)$ be the extended Kalman filter "nominal" state trajectory on the interval $[t_i, t_{i+1})$. Thus, $\mathbf{x}_n\left(t/t_i\right)$ is the solution to Equation (2.19) subject to initial condition $\widehat{\mathbf{x}}\left(t_i^+\right)$, solved on the interval $[t_i, t_{i+1})$:

$$\begin{aligned} \dot{\mathbf{x}}_n\left(t/t_i\right) &= \mathbf{f}\left[\mathbf{x}_n\left(t/t_i\right), \mathbf{u}\left(t_i\right), t\right] \\ &\text{for} \\ \mathbf{x}_n\left(t_i/t_i\right) &= \widehat{\mathbf{x}}\left(t_i^+\right) \end{aligned} \tag{2.27}$$

where $\mathbf{f}\left[\cdot,\cdot,\cdot\right]$ is the dynamics model from Equation (2.17) and $\widehat{\mathbf{x}}\left(t_i^+\right)$ is the state estimate after measurement update at time t_i.

To propagate the state to the next measurement time t_{i+1}, employ the re-linearized $\mathbf{F}\left[t; \mathbf{x}_n\left(t/t_i\right)\right]$ where $\mathbf{F}\left[\cdot\ ;\ \cdot\right]$ is defined in Equation (2.24). Let $\widehat{\boldsymbol{\delta}\mathbf{x}}\left(t/t_i\right)$ denote the estimate of $\boldsymbol{\delta}\mathbf{x}\left(t\right)$ based on measurements through time t_i. Then, on the interval $\left[t_i, t_{i+1}\right)$, $\widehat{\boldsymbol{\delta}\mathbf{x}}\left(t/t_i\right)$ is the solution to

$$\dot{\widehat{\boldsymbol{\delta}\mathbf{x}}}\left(t/t_i\right) = \mathbf{F}\left[t; \mathbf{x}_n\left(t/t_i\right)\right] \widehat{\boldsymbol{\delta}\mathbf{x}}\left(t/t_i\right) \tag{2.28}$$

subject to the initial condition

$$\widehat{\boldsymbol{\delta}\mathbf{x}}\left(t_i/t_i\right) = \widehat{\boldsymbol{\delta}\mathbf{x}}\left(t_i^+\right) = \mathbf{0} \tag{2.29}$$

This initial condition comes from the fact that the best estimate of $\boldsymbol{\delta}\mathbf{x}\left(t_i\right)$ immediately after measurement incorporation is $\mathbf{0}$ due to the re-linearization. Solving Equation (2.28) subject to Equation (2.29) leads to $\widehat{\boldsymbol{\delta}\mathbf{x}}\left(t/t_i\right) = \mathbf{0}$ over the interval $[t_i, t_{i+1})$. Thus, the propagated state perturbation estimate $\widehat{\boldsymbol{\delta}\mathbf{x}}\left(t_{i+1}^-\right)$ is $\mathbf{0}$.

Next, the filter is updated with a measurement as in Equation (2.25), with $\mathbf{z}_n(t_{i+1})$ replaced by

$$\mathbf{z}_n(t_{i+1}/t_i) \stackrel{\triangle}{=} \mathbf{h}[\mathbf{x}_n(t_{i+1}/t_i), t_{i+1}] \tag{2.30}$$

where $\mathbf{z}_n(t_{i+1}/t_i)$ is the measurement prediction based on re-linearization about the most recent state estimate. Given the measurements of (2.25) and (2.30), the measurement update of the state perturbation estimate is

$$\begin{aligned} \widehat{\boldsymbol{\delta}\mathbf{x}}(t_{i+1}^+) &= \widehat{\boldsymbol{\delta}\mathbf{x}}(t_{i+1}^-) + \mathbf{K}(t_{i+1})\left[\{\mathbf{z}_{i+1} - \mathbf{z}_n(t_{i+1}/t_i)\} - \mathbf{H}(t_{i+1})\,\widehat{\boldsymbol{\delta}\mathbf{x}}(t_{i+1}^-)\right] \\ &= \mathbf{K}(t_{i+1})\{\mathbf{z}_{i+1} - \mathbf{h}[\mathbf{x}_n(t_{i+1}/t_i), t_{i+1}]\} \end{aligned} \tag{2.31}$$

where

$\widehat{\boldsymbol{\delta}\mathbf{x}}(t_{i+1}^+)$	is the updated state perturbation estimate
$\widehat{\boldsymbol{\delta}\mathbf{x}}(t_{i+1}^-)$	is the propagated state perturbation estimate ($\widehat{\boldsymbol{\delta}\mathbf{x}}(t_{i+1}^-) = \mathbf{0}$ from the explanation following Equation (2.29))
$\mathbf{K}(t_{i+1})$	is the Kalman filter gain matrix at time t_{i+1}
$\mathbf{z}_{i+1}$	is the measurement realization at time t_{i+1}
$\mathbf{h}[\cdot,\cdot]$	is the general measurement model from Equation (2.17)

Note that $\mathbf{K}(t_{i+1})$ is computed from $\mathbf{P}(t_{i+1}^-)$ and $\mathbf{H}(t_{i+1})$ evaluated along $\mathbf{x}_n(t/t_i)$ on the interval $[t_i, t_{i+1})$. A specific equation for $\mathbf{K}(t_{i+1})$ is given at the end of this subsection.

With the perturbation estimate updated, the next step is to update the whole state estimate. As in the linearized Kalman filter, it is assumed that an adequate model for the whole state is given by

$$\mathbf{x}(t) = \mathbf{x}_n(t/t_i) + \boldsymbol{\delta}\mathbf{x}(t) \tag{2.32}$$

Using Equations (2.31) and (2.32), the optimal estimate of the whole state is defined by

$$\widehat{\mathbf{x}}(t/t_i) \stackrel{\triangle}{=} \mathbf{x}_n(t/t_i) + \widehat{\boldsymbol{\delta}\mathbf{x}}(t/t_i) \tag{2.33}$$

Next, the whole state estimate is propagated to the next measurement time. Since it was previously shown that $\widehat{\boldsymbol{\delta \mathbf{x}}}(t/t_i) = \mathbf{0}$ over the entire interval between measurements, the best state estimate between measurements is given by the solution to Equation (2.27) with appropriate substitutions of state estimates

$$\begin{aligned} \dot{\hat{\mathbf{x}}}(t/t_i) &= \mathbf{f}\left[\hat{\mathbf{x}}(t/t_i), \mathbf{u}(t), t\right] \\ &\text{for} \\ \hat{\mathbf{x}}(t_i/t_i) &= \hat{\mathbf{x}}\left(t_i^+\right) \end{aligned} \tag{2.34}$$

where

$\hat{\mathbf{x}}(t/t_i)$ is the whole state estimate on the interval $[t_i, t_{i+1})$ after measurement incorporation at time t_i

$\mathbf{f}[\cdot,\cdot,\cdot]$ is the general dynamics model

$\hat{\mathbf{x}}\left(t_i^+\right)$ is the whole state estimate immediately following measurement incorporation at time t_i

Finally, the measurement update at time t_{i+1} is incorporated using

$$\begin{aligned} \hat{\mathbf{x}}\left(t_{i+1}^+\right) &= \hat{\mathbf{x}}(t_{i+1}/t_i) + \widehat{\boldsymbol{\delta \mathbf{x}}}\left(t_{i+1}^+\right) \\ &= \hat{\mathbf{x}}(t_{i+1}/t_i) + \mathbf{K}(t_{i+1})\left\{\mathbf{z}_{i+1} - \mathbf{h}\left[\hat{\mathbf{x}}(t_{i+1}/t_i), t_{i+1}\right]\right\} \end{aligned} \tag{2.35}$$

where

$\hat{\mathbf{x}}\left(t_{i+1}^+\right)$ is the updated whole state estimate after measurement incorporation at time t_{i+1}

$\hat{\mathbf{x}}(t_{i+1}/t_i)$ is the propagated whole state estimate from Equation (2.33)

$\mathbf{h}[\cdot,\cdot]$ is the general measurement model

The extended Kalman filter is summarized by the update and propagation algorithm shown in Table 2.2.

Table 2.2 Extended Kalman Filter Summary [30]

Measurement Update

$$\begin{aligned}
\mathbf{K}(t_i) &= \mathbf{P}(t_i^-)\,\mathbf{H}^T\left[t_i;\widehat{\mathbf{x}}(t_i^-)\right] \\
&\quad \{\,\mathbf{H}\left[t_i;\widehat{\mathbf{x}}(t_i^-)\right]\mathbf{P}(t_i^-)\,\mathbf{H}^T\left[t_i;\widehat{\mathbf{x}}(t_i^-)\right]+\mathbf{R}(t_i)\,\}^{-1} \\
\widehat{\mathbf{x}}(t_i^+) &= \widehat{\mathbf{x}}(t_i^-) + \mathbf{K}(t_i)\left\{\mathbf{z}_i - \mathbf{h}\left[\widehat{\mathbf{x}}(t_i^-),t_i\right]\right\} \\
\mathbf{P}(t_i^+) &= \mathbf{P}(t_i^-) - \mathbf{K}(t_i)\,\mathbf{H}\left[t_i;\widehat{\mathbf{x}}(t_i^-)\right]\mathbf{P}(t_i^-)
\end{aligned}$$

where

$$\mathbf{H}\left[t_i;\widehat{\mathbf{x}}(t_i^-)\right] \triangleq \left.\frac{\partial\mathbf{h}[\mathbf{x},t_i]}{\partial\mathbf{x}}\right|_{\mathbf{x}=\widehat{\mathbf{x}}(t_i^-)}$$

State Propagation

Integrate over $t_i \le t \le t_{i+1}$

$$\begin{aligned}
\dot{\widehat{\mathbf{x}}}(t/t_i) &= \mathbf{f}\left[\widehat{\mathbf{x}}(t/t_i),\mathbf{u}(t),t\right] \\
\dot{\mathbf{P}}(t/t_i) &= \mathbf{F}\left[t;\widehat{\mathbf{x}}(t/t_i)\right]\mathbf{P}(t/t_i) + \mathbf{P}(t/t_i)\,\mathbf{F}^T\left[t;\widehat{\mathbf{x}}(t/t_i)\right] \\
&\quad +\mathbf{G}(t)\,\mathbf{Q}(t)\,\mathbf{G}^T(t)
\end{aligned}$$

subject to

$$\begin{aligned}
\widehat{\mathbf{x}}(t_i/t_i) &= \widehat{\mathbf{x}}(t_i^+) \\
\mathbf{P}(t_i/t_i) &= \mathbf{P}(t_i^+)
\end{aligned}$$

where

$$\mathbf{F}\left[t;\widehat{\mathbf{x}}(t/t_i)\right] \triangleq \left.\frac{\partial\mathbf{f}[\mathbf{x},\,\mathbf{u}(t),t]}{\partial\mathbf{x}}\right|_{\mathbf{x}=\widehat{\mathbf{x}}(t/t_i)}$$

After Integration Define

$$\begin{aligned}
\widehat{\mathbf{x}}(t_{i+1}^-) &\triangleq \widehat{\mathbf{x}}(t_{i+1}/t_i) \\
\mathbf{P}(t_{i+1}^-) &\triangleq \mathbf{P}(t_{i+1}/t_i)
\end{aligned}$$

2.3 Multiple Model Adaptive Estimator

The topic of the Multiple Model Adaptive Estimator is introduced in this section. The MMAE was initially proposed for GPS carrier-phase integer ambiguity resolution by [21] with a modified version tested by [18]. The idea of using the MMAE (then termed a Magill filter) for GPS carrier-phase integer ambiguity resolution was suggested by [21] and simulations were performed showing the technique had potential for application in static scenarios. The technique was later modified slightly by [18], and applied to kinematic GPS data, further demonstrating the technique's potential. This section provides background on the traditional MMAE technique. A later section will describe more fully the approaches of [18, 21].

Multiple Model Adaptive Estimation was originally proposed in [26]. The technique has since been described in a number of textbooks on Kalman filtering. The presentation here is similar to that found in [30].

Consider that the model of Equation (2.1) contains an unknown parameter which varies slowly, if at all, when compared to the states. Further, assume that the parameter can take on values that are discrete members of a finite set $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_K\}$. To achieve the best state estimation performance, the unknown parameter must be estimated. One approach is to use the MMAE shown in Figure 1.2. The MMAE is a bank of K linear Kalman filters, each built around a single member of the parameter set $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_K\}$. Each individual filter is referred to as an elemental filter. The elemental filter outputs are combined in an optimally weighted fashion to produce the state vector estimate. Additionally, the optimal weights can be used to determine the best estimate of the correct parameter $\mathbf{a}_k$, where $k \in \{1, 2, \ldots, K\}$. With this introduction in mind, the MMAE algorithm is derived.

Let $\mathbf{a}$ represent an uncertain parameter in the system model of Equation (2.1). In general, $\mathbf{a}$ may affect any or all of the model matrices: $\mathbf{\Phi}$, $\mathbf{B}_d$, $\mathbf{H}$, $\mathbf{Q}_d$, or $\mathbf{R}$. A goal of the MMAE is to find the mean of the conditional PDF given on the left hand side of

$$f_{\mathbf{x}(t_i),\mathbf{a}|\mathbf{Z}(t_i)}\left(\boldsymbol{\xi}, \boldsymbol{\alpha}|\mathbf{Z}\left(t_i\right) = \mathbf{Z}_i\right) = f_{\mathbf{x}(t_i)|\mathbf{a},\mathbf{Z}(t_i)}\left(\boldsymbol{\xi}|\mathbf{a} = \boldsymbol{\alpha}, \mathbf{Z}\left(t_i\right) = \mathbf{Z}_i\right) f_{\mathbf{a}|\mathbf{Z}(t_i)}\left(\boldsymbol{\alpha}|\mathbf{Z}\left(t_i\right) = \mathbf{Z}_i\right) \tag{2.36}$$

where $f_{\mathbf{x}(t_i),\mathbf{a}|\mathbf{Z}(t_i)}\left(\boldsymbol{\xi}, \boldsymbol{\alpha}|\mathbf{Z}\left(t_i\right) = \mathbf{Z}_i\right)$ is the joint conditional PDF for the state vector and parameter set, $\mathbf{x}\left(t_i\right)$ and $\mathbf{a}$, conditioned on the realized measurement history $\mathbf{Z}_i$.

Equation (2.36) is a statement of Baye's rule for this conditional PDF. In defining the desired probability density, consider the first density on the right hand side of Equation (2.36) $f_{\mathbf{x}(t_i)|\mathbf{a},\mathbf{Z}(t_i)}\left(\boldsymbol{\xi}|\mathbf{a} = \boldsymbol{\alpha}, \mathbf{Z}\left(t_i\right) = \mathbf{Z}_i\right)$. Since this density is conditioned on a particular realization of $\mathbf{a}$, it is analogous to the standard conditional density used in Kalman filtering for state estimation. Therefore, under assumptions of Equations (2.1) through (2.3), it is a Gaussian density with mean $\widehat{\mathbf{x}}_k\left(t_i^+\right)$ and covariance $\mathbf{P}_k\left(t_i^+\right)$ as calculated by a standard Kalman filter for a given realization $\mathbf{a}_k$ of the parameter $\mathbf{a}$ (where the subscript k indicates the dependence upon the realized value of the parameter $\mathbf{a}$).

Next, consider the second density on the right hand side of Equation (2.36). Although it has already been assumed that the elements of $\mathbf{a}$ take on discrete values, consider for a moment the case in which $\mathbf{a}$ is a continuous (rather than discrete) random variable. Using Baye's rule, the second density on the right hand side of Equation (2.36) can be expressed as

$$f_{\mathbf{a}|\mathbf{Z}(t_i)}\left(\boldsymbol{\alpha}|\mathbf{Z}\left(t_i\right)=\boldsymbol{\mathcal{Z}}_i\right) = \frac{f_{\mathbf{z}(t_i)|\mathbf{a},\mathbf{Z}(t_{i-1})}\left(\boldsymbol{\zeta}|\mathbf{a}=\boldsymbol{\alpha},\mathbf{Z}\left(t_{i-1}\right)=\boldsymbol{\mathcal{Z}}_{i-1}\right) f_{\mathbf{a}|\mathbf{Z}(t_{i-1})}\left(\boldsymbol{\alpha}|\mathbf{Z}\left(t_{i-1}\right)=\boldsymbol{\mathcal{Z}}_{i-1}\right)}{\int\limits_{A} f_{\mathbf{z}(t_i)|\mathbf{a},\mathbf{Z}(t_{i-1})}\left(\boldsymbol{\zeta}|\mathbf{a}=\boldsymbol{\alpha},\mathbf{Z}\left(t_{i-1}\right)=\boldsymbol{\mathcal{Z}}_{i-1}\right) f_{\mathbf{a}|\mathbf{Z}(t_{i-1})}\left(\boldsymbol{\alpha}|\mathbf{Z}\left(t_{i-1}\right)=\boldsymbol{\mathcal{Z}}_{i-1}\right)\mathbf{d}\boldsymbol{\alpha}} \tag{2.37}$$

From the system model and assumptions in Equations (2.1) to (2.3), the density $f_{\mathbf{z}(t_i)|\mathbf{a},\mathbf{Z}(t_{i-1})}\left(\boldsymbol{\zeta}_i|\mathbf{a}=\boldsymbol{\alpha},\mathbf{Z}\left(t_{i-1}\right)=\boldsymbol{\mathcal{Z}}_{i-1}\right)$ is Gaussian [29]. Its mean is the measurement prediction $\mathbf{H}\left(t_i\right)\widehat{\mathbf{x}}\left(t_i^-\right)$, and its covariance is $\left[\mathbf{H}\left(t_i\right)\mathbf{P}\left(t_i^-\right)\mathbf{H}^T\left(t_i\right)+\mathbf{R}\left(t_i\right)\right]$, as computed in a Kalman filter based on the parameter vector value $\boldsymbol{\alpha}$. Given an a priori density $f_{\mathbf{a}}\left(\boldsymbol{\alpha}\right)$, a state estimate is given by the conditional mean $E\left\{\mathbf{x}\left(t_i\right)|\mathbf{Z}\left(t_i\right)=\mathbf{Z}_i\right\}$.

Now, consider the case of $\mathbf{a}$ taking on discrete values. The a priori density for $\mathbf{a}$ can be defined by

$$f_{\mathbf{a}(t_0)}\left(\boldsymbol{\alpha}\right) = \sum_{k=1}^{K} p_k\left(t_0\right)\delta\left(\boldsymbol{\alpha} - \mathbf{a}_k\right) \tag{2.38}$$

In Equation (2.38), $p_k(t_0)$ is the probability that $\mathbf{a}$ takes on discrete value $\mathbf{a}_k$ at initial time t_0, and $\delta\left(\boldsymbol{\alpha} - \mathbf{a}_k\right)$ is the Dirac delta function. Thus, $f_{\mathbf{a}}\left(\boldsymbol{\alpha}\right)$ is a summation of Dirac delta functions weighted by appropriate probabilities, forming a discrete PDF. In order for Equation (2.38) to be a valid PDF, it must meet the conditions

$$\begin{aligned} p_k\left(t_0\right) &\geq 0 \quad \forall\ k \\ \sum_{k=1}^{K} p_k\left(t_0\right) &= 1 \end{aligned} \tag{2.39}$$

Note that the values of $p_k(t_0)$ in Equation (2.39) represent the initial probabilities of each element of the set $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_K\}$. In the absence of scientific knowledge or experimental data, the parameter hypotheses are initially assumed to be equally likely, leading to $p_k(t_0) = \frac{1}{K} \ \forall\, k$.

Next, consider the measurement effects, and define the hypothesis conditional probability

$$p_k(t_i) \triangleq \text{prob}\{\mathbf{a} = \mathbf{a}_k | \mathbf{Z}(t_i) = \mathbf{Z}_i\} \tag{2.40}$$

Thus, $p_k(t_i)$ is the conditional probability that $\mathbf{a} = \mathbf{a}_k$, given the realized measurement history $\mathbf{Z}_i$.

Substituting Equations (2.38) and (2.40) into Equation (2.37), and using the sifting property of the Dirac delta function, yields

$$p_k(t_i) = \frac{f_{\mathbf{z}(t_i)|\mathbf{a},\mathbf{Z}(t_{i-1})}(\mathbf{z}_i|\mathbf{a} = \mathbf{a}_k, \mathbf{Z}(t_{i-1}) = \mathbf{Z}_{i-1})\, p_k(t_{i-1})}{\sum_{j=1}^{K} f_{\mathbf{z}(t_i)|\mathbf{a},\mathbf{Z}(t_{i-1})}(\mathbf{z}_i|\mathbf{a} = \mathbf{a}_j, \mathbf{Z}(t_{i-1}) = \mathbf{Z}_{i-1})\, p_j(t_{i-1})} \tag{2.41}$$

Recall from the definition of Equation (2.40) that $p_k(t_i)$ is the conditional probability for the parameter $\mathbf{a}_k$ given the measurement history $\mathbf{Z}_i$. Thus, if desired, Equation (2.41) is used to calculate the best estimate of the parameter. The formula for the optimal parameter estimate will be shown later in Equation (2.48).

Now, consider state estimation. The optimal state estimate is derived as follows [30]

$$\begin{aligned}
\widehat{\mathbf{x}}(t_i^+) &= E\{\mathbf{x}(t_i)|\mathbf{Z}(t_i) = \mathbf{Z}_i\} \\
&= \int_{-\infty}^{\infty} \boldsymbol{\xi} \left[\int_A f_{\mathbf{x}(t_i),\mathbf{a}|\mathbf{Z}(t_i)}(\boldsymbol{\xi}, \boldsymbol{\alpha}|\mathbf{Z}(t_i) = \mathbf{Z}_i)\, d\boldsymbol{\alpha} \right] d\boldsymbol{\xi} \\
&= \int_{-\infty}^{\infty} \boldsymbol{\xi} \left[\int_A f_{\mathbf{x}(t_i)|\mathbf{a},\mathbf{Z}(t_i)}(\boldsymbol{\xi}|\mathbf{a} = \boldsymbol{\alpha}, \mathbf{Z}(t_i) = \mathbf{Z}_i)\, f_{\mathbf{a}|\mathbf{Z}(t_i)}(\boldsymbol{\alpha}|\mathbf{Z}(t_i) = \mathbf{Z_i})\, d\boldsymbol{\alpha} \right] d\boldsymbol{\xi} \\
&= \int_{-\infty}^{\infty} \boldsymbol{\xi} \left[\int_A f_{\mathbf{x}(t_i)|\mathbf{a},\mathbf{Z}(t_i)}(\boldsymbol{\xi}|\mathbf{a} = \boldsymbol{\alpha}, \mathbf{Z}(t_i) = \mathbf{Z}_i) \sum_{k=1}^{K} p_k(t_i)\, \delta(\boldsymbol{\alpha} - \mathbf{a}_k)\, d\boldsymbol{\alpha} \right] d\boldsymbol{\xi}
\end{aligned}$$

$$\widehat{\mathbf{x}}\left(t_i^+\right) \;=\; \int_{-\boldsymbol{\infty}}^{\boldsymbol{\infty}} \boldsymbol{\xi} \left[\sum_{k=1}^{K} f_{\mathbf{x}(t_i)|\mathbf{a},\mathbf{Z}(t_i)}\left(\boldsymbol{\xi}|\mathbf{a}=\mathbf{a}_k, \mathbf{Z}\left(t_i\right)=\mathbf{Z}_i\right) p_k\left(t_i\right)\right] \boldsymbol{d\xi} \tag{2.42}$$

where Equation (2.36) was applied in the third line of Equation (2.42). In the fourth and fifth lines of Equation (2.42), $f_{\mathbf{a}|\mathbf{Z}(t_i)}\left(\boldsymbol{\alpha}|\mathbf{Z}\left(t_i\right)=\mathbf{Z_i}\right)$ is replaced by $\sum_{k\,=\,1}^{K}\, p_k\left(t_i\right)\delta\left(\boldsymbol{\alpha}\;-\;\mathbf{a}_k\right)$ by invoking the definition of $p_k\left(t_i\right)$. Also, rather than writing a series of integrals over each element of the vector $\boldsymbol{\xi}$, the common engineering notation of writing a single integral with boldface limits $\left(\int_{-\boldsymbol{\infty}}^{\boldsymbol{\infty}} \{\cdot\}\, \boldsymbol{d\xi}\right)$ was used. Assuming $\boldsymbol{\xi}$ is n-dimensional, this notation is interpreted as

$$\int_{-\boldsymbol{\infty}}^{\boldsymbol{\infty}} \{\cdot\}\, \boldsymbol{d\xi} \;=\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \{\cdot\}\, d\xi_1 d\xi_2 \ldots d\xi_n$$

Since the summation in Equation (2.42) is finite, the order of integration and summation can be interchanged [1] yielding

$$\widehat{\mathbf{x}}\left(t_i^+\right) \;=\; \sum_{k=1}^{K}\left[\int_{-\boldsymbol{\infty}}^{\boldsymbol{\infty}} \boldsymbol{\xi} f_{\mathbf{x}(t_i)|\mathbf{a},\mathbf{Z}(t_i)}\left(\boldsymbol{\xi}|\mathbf{a}=\mathbf{a_k}, \mathbf{Z}\left(t_i\right)=\mathbf{Z_i}\right) \boldsymbol{d\xi}\right] p_k\left(t_i\right) \tag{2.43}$$

The term inside the square bracket of Equation (2.43) is simply the optimal state estimate conditioned upon a given parameter value $\mathbf{a}_k$ and the realized measurement history $\mathbf{Z}_i$. Define this quantity to be $\widehat{\mathbf{x}}_k\left(t_i^+\right)$. The updated state estimate then takes on its final form given by

$$\widehat{\mathbf{x}}\left(t_i^+\right) \;=\; \sum_{k=1}^{K} \widehat{\mathbf{x}}_k\left(t_i^+\right) p_k\left(t_i\right) \tag{2.44}$$

Equation (2.44) shows that the final updated state estimate is the weighted sum of the elemental filter state estimates, weighted by the probability that the parameter for that particular elemental filter is the correct parameter (based on the measurements seen to date). This is exactly the form that is shown in Figure 1.2.

Next, consider the covariance update for the MMAE. From the definition, the covariance is

$$\begin{aligned}
\mathbf{P}\left(t_i^+\right) &= E\left\{\left[\mathbf{x}\left(t_i\right)-\widehat{\mathbf{x}}\left(t_i^+\right)\right]\left[\mathbf{x}\left(t_i\right)-\widehat{\mathbf{x}}\left(t_i^+\right)\right]^T|\mathbf{Z}\left(t_i\right)=\mathbf{Z}_i\right\} \\
&= \int_{-\infty}^{\infty}\left[\boldsymbol{\xi}-\widehat{\mathbf{x}}\left(t_i^+\right)\right]\left[\boldsymbol{\xi}-\widehat{\mathbf{x}}\left(t_i^+\right)\right]^T f_{\mathbf{x}(t_i)|\mathbf{Z}(t_i)}\left(\boldsymbol{\xi}|\mathbf{Z}\left(t_i\right)=\mathbf{Z}_i\right)d\boldsymbol{\xi} \\
&= \int_{-\infty}^{\infty}\left[\boldsymbol{\xi}-\widehat{\mathbf{x}}\left(t_i^+\right)\right]\left[\boldsymbol{\xi}-\widehat{\mathbf{x}}\left(t_i^+\right)\right]^T \int_A f_{\mathbf{x}(t_i),\mathbf{a}|\mathbf{Z}(t_i)}\left(\boldsymbol{\xi},\boldsymbol{\alpha}|\mathbf{Z}\left(t_i\right)=\mathbf{Z}_i\right)d\boldsymbol{\alpha}\;d\boldsymbol{\xi} \\
&= \int_{-\infty}^{\infty}\left[\boldsymbol{\xi}-\widehat{\mathbf{x}}\left(t_i^+\right)\right]\left[\boldsymbol{\xi}-\widehat{\mathbf{x}}\left(t_i^+\right)\right]^T \\
&\quad \int_A f_{\mathbf{x}(t_i)|\mathbf{a},\mathbf{Z}(t_i)}\left(\boldsymbol{\xi}|\mathbf{a}=\boldsymbol{\alpha},\mathbf{Z}\left(t_i\right)=\mathbf{Z}_i\right) f_{\mathbf{a}|\mathbf{Z}(t_i)}\left(\boldsymbol{\alpha}|\mathbf{Z}\left(t_i\right)=\mathbf{Z}_i\right)d\boldsymbol{\alpha}\;d\boldsymbol{\xi} \\
\mathbf{P}\left(t_i^+\right) &= \int_{-\infty}^{\infty}\left[\boldsymbol{\xi}-\widehat{\mathbf{x}}\left(t_i^+\right)\right]\left[\boldsymbol{\xi}-\widehat{\mathbf{x}}\left(t_i^+\right)\right]^T \sum_{k=1}^{K} f_{\mathbf{x}(t_i)|\mathbf{a},\mathbf{Z}(t_i)}\left(\boldsymbol{\xi}|\mathbf{a}=\mathbf{a}_k,\mathbf{Z}\left(t_i\right)=\mathbf{Z}_i\right)p_k\left(t_i\right)\;d\boldsymbol{\xi}
\end{aligned} \tag{2.45}$$

Again, since the summation is finite, the order of integration and summation can be interchanged. Additionally, $f_{\mathbf{a}|\mathbf{Z}(t_i)}\left(\boldsymbol{\alpha}|\mathbf{Z}\left(t_i\right)=\mathbf{Z}_i\right)$ may be brought outside the integral in the form $p_k\left(t_i\right)$ using Equation (2.40). Equation (2.45) then becomes

$$\begin{aligned}
\mathbf{P}\left(t_i^+\right) &= \sum_{k=1}^{K} p_k\left(t_i\right)\left\{\int_{-\infty}^{\infty}\left[\boldsymbol{\xi}-\widehat{\mathbf{x}}\left(t_i^+\right)\right]\left[\boldsymbol{\xi}-\widehat{\mathbf{x}}\left(t_i^+\right)\right]^T f_{\mathbf{x}(t_i)|\mathbf{a},\mathbf{Z}(t_i)}\left(\boldsymbol{\xi}|\mathbf{a}=\mathbf{a}_k,\mathbf{Z}\left(t_i\right)=\mathbf{Z}_i\right)d\boldsymbol{\xi}\right\} \\
&= \sum_{k=1}^{K} p_k\left(t_i\right)\left\{\int_{-\infty}^{\infty}\boldsymbol{\xi}\boldsymbol{\xi}^T f_{\mathbf{x}(t_i)|\mathbf{a},\mathbf{Z}(t_i)}\left(\boldsymbol{\xi}|\mathbf{a}=\mathbf{a}_k,\mathbf{Z}\left(t_i\right)=\mathbf{Z}_i\right)d\boldsymbol{\xi}\right. \\
&\quad -\int_{-\infty}^{\infty}\widehat{\mathbf{x}}\left(t_i^+\right)\boldsymbol{\xi}^T f_{\mathbf{x}(t_i)|\mathbf{a},\mathbf{Z}(t_i)}\left(\boldsymbol{\xi}|\mathbf{a}=\mathbf{a}_k,\mathbf{Z}\left(t_i\right)=\mathbf{Z}_i\right)d\boldsymbol{\xi} \\
&\quad -\int_{-\infty}^{\infty}\boldsymbol{\xi}\widehat{\mathbf{x}}^T\left(t_i^+\right) f_{\mathbf{x}(t_i)|\mathbf{a},\mathbf{Z}(t_i)}\left(\boldsymbol{\xi}|\mathbf{a}=\mathbf{a}_k,\mathbf{Z}\left(t_i\right)=\mathbf{Z}_i\right)d\boldsymbol{\xi} \\
&\quad \left.+\int_{-\infty}^{\infty}\widehat{\mathbf{x}}\left(t_i^+\right)\widehat{\mathbf{x}}^T\left(t_i^+\right) f_{\mathbf{x}(t_i)|\mathbf{a},\mathbf{Z}(t_i)}\left(\boldsymbol{\xi}|\mathbf{a}=\mathbf{a}_k,\mathbf{Z}\left(t_i\right)=\mathbf{Z}_i\right)d\boldsymbol{\xi}\right\}
\end{aligned}$$

$$
\begin{aligned}
\mathbf{P}\left(t_i^+\right) &= \sum_{k=1}^{K} p_k\left(t_i\right) \left\{ \int_{-\infty}^{\infty} \boldsymbol{\xi}\boldsymbol{\xi}^T f_{\mathbf{x}(t_i)|\mathbf{a},\mathbf{Z}(t_i)}\left(\boldsymbol{\xi}|\mathbf{a}=\mathbf{a}_k, \mathbf{Z}\left(t_i\right)=\mathbf{Z}_i\right) d\boldsymbol{\xi} \right. \\
&\quad - \widehat{\mathbf{x}}\left(t_i^+\right) \int_{-\infty}^{\infty} \boldsymbol{\xi}^T f_{\mathbf{x}(t_i)|\mathbf{a},\mathbf{Z}(t_i)}\left(\boldsymbol{\xi}|\mathbf{a}=\mathbf{a}_k, \mathbf{Z}\left(t_i\right)=\mathbf{Z}_i\right) d\boldsymbol{\xi} \\
&\quad - \left[\int_{-\infty}^{\infty} \boldsymbol{\xi} f_{\mathbf{x}(t_i)|\mathbf{a},\mathbf{Z}(t_i)}\left(\boldsymbol{\xi}|\mathbf{a}=\mathbf{a}_k, \mathbf{Z}\left(t_i\right)=\mathbf{Z}_i\right) d\boldsymbol{\xi} \right] \widehat{\mathbf{x}}^T\left(t_i^+\right) \\
&\quad \left. + \widehat{\mathbf{x}}\left(t_i^+\right) \widehat{\mathbf{x}}^T\left(t_i^+\right) \int_{-\infty}^{\infty} f_{\mathbf{x}(t_i)|\mathbf{a},\mathbf{Z}(t_i)}\left(\boldsymbol{\xi}|\mathbf{a}=\mathbf{a}_k, \mathbf{Z}\left(t_i\right)=\mathbf{Z}_i\right) d\boldsymbol{\xi} \right\} \\
&= \sum_{k=1}^{K} p_k\left(t_i\right) \left\{ \int_{-\infty}^{\infty} \boldsymbol{\xi}\boldsymbol{\xi}^T f_{\mathbf{x}(t_i)|\mathbf{a},\mathbf{Z}(t_i)}\left(\boldsymbol{\xi}|\mathbf{a}=\mathbf{a}_k, \mathbf{Z}\left(t_i\right)=\mathbf{Z}_i\right) d\boldsymbol{\xi} - \widehat{\mathbf{x}}\left(t_i^+\right) \widehat{\mathbf{x}}_k^T\left(t_i^+\right) \right. \\
&\quad \left. - \widehat{\mathbf{x}}_k\left(t_i^+\right) \widehat{\mathbf{x}}^T\left(t_i^+\right) + \widehat{\mathbf{x}}\left(t_i^+\right) \widehat{\mathbf{x}}^T\left(t_i^+\right) \right\} \\
\mathbf{P}\left(t_i^+\right) &= \sum_{k=1}^{K} p_k\left(t_i\right) \left\{ \mathbf{P}_k\left(t_i^+\right) + \left[\widehat{\mathbf{x}}_k\left(t_i^+\right) - \widehat{\mathbf{x}}\left(t_i^+\right)\right] \left[\widehat{\mathbf{x}}_k\left(t_i^+\right) - \widehat{\mathbf{x}}\left(t_i^+\right)\right]^T \right\} \qquad (2.46)
\end{aligned}
$$

Thus, the state estimate error covariance of Equation (2.46) is dependent upon the overall state estimate $\widehat{\mathbf{x}}\left(t_i^+\right)$, the state estimate from each elemental filter $\widehat{\mathbf{x}}_k\left(t_i^+\right)$, the error covariance from each elemental filter $\mathbf{P}_k\left(t_i^+\right)$, and the probability $p_k\left(t_i\right)$ that each elemental filter contains the correct parameter value. The estimates $\widehat{\mathbf{x}}_k\left(t_i^+\right)$ and $\mathbf{P}_k\left(t_i^+\right)$ are readily available from each elemental Kalman filter. If $p_k\left(t_i\right)$ is available, then the filter shown in Figure 1.2 is defined. Note that $\mathbf{P}\left(t_i^+\right)$ is not required for an online MMAE. As a final step, consider the calculation of $p_k\left(t_i\right)$.

Given an initial estimate $p_k\left(t_0\right)$ subject to conditions of Equation (2.39), the probabilities $p_k\left(t_i\right)$ can be updated at each measurement epoch using Equation (2.41). The only value still needed for the update is the density $f_{\mathbf{z}(t_i)|\mathbf{a},\mathbf{Z}(t_{i-1})}\left(\mathbf{z}_{\boldsymbol{i}}|\mathbf{a}=\mathbf{a}_k, \mathbf{Z}\left(t_{i-1}\right)=\mathbf{Z}_{i-1}\right)$ for $k = 1, 2, \ldots K$, which is calculated as [30]

$$f_{\mathbf{z}(t_i)|\mathbf{a},\mathbf{Z}(t_{i-1})}\left(\mathbf{z}_i|\mathbf{a}=\mathbf{a}_k,\mathbf{Z}\left(t_{i-1}\right)=\mathbf{Z}_{i-1}\right)$$
$$= \frac{1}{(2\pi)^{\frac{m}{2}}\left|\mathbf{A}_k\left(t_i^-\right)\right|^{\frac{1}{2}}}\exp\left\{-\frac{1}{2}\mathbf{r}_k^T\left(t_i^-\right)\mathbf{A}_k^{-1}\left(t_i^-\right)\mathbf{r}_k\left(t_i^-\right)\right\} \tag{2.47}$$

where $\mathbf{A}_k\left(t_i^-\right) = \mathbf{H}_k\left(t_i\right)\mathbf{P}_k\left(t_i^-\right)\mathbf{H}_k^T\left(t_i\right) + \mathbf{R}\left(t_i\right)$ is the covariance matrix of the measurement residual vector $\mathbf{r}_k\left(t_i^-\right)$ from elemental filter k. Both $\mathbf{r}_k\left(t_i^-\right)$ and $\mathbf{A}_k\left(t_i^-\right)$ are calculated from each of the K elemental filters, and $\mathbf{r}_k\left(t_i^-\right) = \mathbf{z}_i - \mathbf{H}\left(t_i\right)\widehat{\mathbf{x}}_k\left(t_i^-\right)$, where $\mathbf{z}_i$ is the realized measurement vector at t_i.

If desired, the best parameter estimate can be derived as follows:

$$\begin{aligned}
\widehat{\mathbf{a}}\left(t_i\right) &\stackrel{\triangle}{=} E\left\{\mathbf{a}\left(t_i\right)|\mathbf{Z}\left(t_i\right)=\mathbf{Z}_i\right\} \\
&= \int_{-\infty}^{\infty}\boldsymbol{\alpha} f_{\mathbf{a}|\mathbf{Z}(t_i)}\left(\boldsymbol{\alpha}|\mathbf{Z}_i\right)d\boldsymbol{\alpha} \\
&= \int_{-\infty}^{\infty}\boldsymbol{\alpha}\left[\sum_{k=1}^{K}p_k\left(t_i\right)\delta\left(\boldsymbol{\alpha}-\mathbf{a}_k\right)\right]d\boldsymbol{\alpha} \\
\widehat{\mathbf{a}}\left(t_i\right) &= \sum_{k=1}^{K}\mathbf{a}_k p_k\left(t_i\right)
\end{aligned} \tag{2.48}$$

As in previous derivations, the last line of Equation (2.48) is the result of interchanging the order of integration and summation, and applying the sifting property of the Dirac delta function.

Finally, the parameter estimate precision is given by the parameter conditional covariance

$$\begin{aligned}
\mathbf{P}_{\mathbf{a}}\left(t_i\right) &= E\left\{\left[\mathbf{a}-\widehat{\mathbf{a}}\left(t_i\right)\right]\left[\mathbf{a}-\widehat{\mathbf{a}}\left(t_i\right)\right]^T|\mathbf{Z}\left(t_i\right)=\mathbf{Z}_i\right\} \\
&= E\left[\mathbf{a}\mathbf{a}^T|\mathbf{Z}\left(t_i\right)=\mathbf{Z}_i\right] - \widehat{\mathbf{a}}\left(t_i\right)E\left[\mathbf{a}^T|\mathbf{Z}\left(t_i\right)=\mathbf{Z}_i\right] \\
&\quad - E\left[\mathbf{a}|\mathbf{Z}\left(t_i\right)=\mathbf{Z}_i\right]\widehat{\mathbf{a}}^T\left(t_i\right) + \widehat{\mathbf{a}}\left(t_i\right)\widehat{\mathbf{a}}^T\left(t_i\right)
\end{aligned}$$

$$
\begin{aligned}
\mathbf{P}_{\mathbf{a}}(t_i) &= \sum_{k=1}^{K} \mathbf{a}_k \mathbf{a}_k^T p_k(t_i) - \widehat{\mathbf{a}}(t_i) \sum_{k=1}^{K} \mathbf{a}_k^T p_k(t_i) \\
&\quad - \sum_{k=1}^{K} \mathbf{a}_k p_k(t_i) \widehat{\mathbf{a}}^T(t_i) + \widehat{\mathbf{a}}(t_i) \widehat{\mathbf{a}}^T(t_i) \sum_{k=1}^{K} p_k(t_i) \\
&= \sum_{k=1}^{K} \left[\mathbf{a}_k \mathbf{a}_k^T - \widehat{\mathbf{a}}(t_i) \mathbf{a}_k^T - \mathbf{a}_k \widehat{\mathbf{a}}^T(t_i) + \widehat{\mathbf{a}}(t_i) \widehat{\mathbf{a}}^T(t_i) \right] p_k(t_i) \\
\mathbf{P}_{\mathbf{a}}(t_i) &= \sum_{k=1}^{K} \left[\mathbf{a}_k - \widehat{\mathbf{a}}^T(t_i) \right] \left[\mathbf{a}_k - \widehat{\mathbf{a}}^T(t_i) \right]^T p_k(t_i) \qquad (2.49)
\end{aligned}
$$

Note that Equations (2.48) and (2.49) need not be calculated to estimate the state $\widehat{\mathbf{x}}\left(t_i^+\right)$. Also, care should be taken when tuning the MMAE [30]. Dynamics pseudonoise, added to account for modeling inadequacies, has a tendency to make measurement residuals from elemental filters have similar magnitudes. However, the MMAE relies on the weighted sum of squared residuals (also called the likelihood quotient) $\mathbf{r}_k^T\left(t_i^-\right) \mathbf{A}_k^{-1}\left(t_i^-\right) \mathbf{r}_k\left(t_i^-\right)$ being substantially different between the elemental filter based on the correct hypothesis and any other elemental filter. Excessive amounts of dynamics pseudonoise being added can incapacitate the adaptivity of the MMAE. So, "conservative" tuning (adding sufficient pseudonoise to the model to preclude filter divergence) is to be avoided. Instead, "tight" filter tuning without much (if any) added pseudonoise is preferred. Therefore, to prevent the MMAE from overemphasizing incorrect elemental filters, care must be taken when tuning the MMAE.

Finally, in this derivation, it was assumed that the correct parameter is a member of a discrete set of parameters (i.e., $\mathbf{a} \in \{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_K\}$). When this assumption is met, the MMAE will converge to the correct parameter value $\mathbf{a}_k$. However, if this assumption is not met, the MMAE will converge to the parameter value which is closest to the correct value in the Baram distance measure sense [5, 6].

2.4 MMAE Performance Enhancements

Several modifications to the traditional MMAE have been proposed to enhance performance. Most are *ad hoc* methods which are devised to address problems in specific

applications. Good summaries of these enhancements are found in [31, 46]. Some of these enhancements are listed and discussed next.

1. β Stripping
2. Probability Lower Bounding
3. Kalman Filter Tuning
4. Elemental Filter Gain Modulation
5. Scalar Penalty Increase
6. Probability Smoothing
7. Increased Residual Propagation
8. Interacting Multiple Model

2.4.1 β Stripping. While using the MMAE for sensor failure identification, [32, 45] found that the MMAE was susceptible to false alarms when the likelihood quotients from all the elemental filters had similar magnitudes over time. Under this condition, the elemental filters are all equally adequate, and the MMAE should equally distribute the hypothesis conditional probabilities among the filters. However, the MMAE instead assigns more probability weight to the elemental filter with the smallest $|\mathbf{A}_k\left(t_i^-\right)|$. This is the so-called "β-dominance" effect [2, 3, 32] caused when

$$\beta_k\left(t_i\right) \;=\; \frac{1}{(2\pi)^{m/2}|\mathbf{A}_k\left(t_i^-\right)|^{1/2}}$$

dominates over

$$\exp\{-\frac{1}{2}\mathbf{r}_k^T\left(t_i^-\right)\mathbf{A}_k^{-1}\left(t_i^-\right)\mathbf{r}_k\left(t_i^-\right)\}$$

in the hypothesis conditional probability calculation. Such behavior by the MMAE is inappropriate and undesirable because the filter-computed residual covariance $\mathbf{A}_k\left(t_i^-\right)$ is completely divorced from the actual sensor measurements. In fact, if linear Kalman filters are used for the elemental filters, then $\mathbf{A}_k\left(t_i^-\right)$ is pre-computable before any measurements are taken.

To prevent the β-dominance effect, [32, 45] stripped the $\beta_k(t_i)$ terms from all the Gaussian densities in Equation (2.41). Thus,

$$\frac{1}{(2\pi)^{\frac{m}{2}} |\mathbf{A}_k(t_i^-)|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}\mathbf{r}_k^T(t_i^-)\mathbf{A}_k^{-1}(t_i^-)\mathbf{r}_k(t_i^-)\right\}$$

was replaced by

$$\exp\left\{-\frac{1}{2}\mathbf{r}_k^T(t_i^-)\mathbf{A}_k^{-1}(t_i^-)\mathbf{r}_k(t_i^-)\right\}$$

in the Gaussian densities of Equation (2.41). While these stripped "densities" do not integrate to 1, and therefore are not true probability densities, the modified hypothesis conditional probability formula is still valid because the denominator still scales the probabilities associated with each elemental filter such that they all sum to 1.

As mentioned at the beginning of this section, sensor failure identification is an application in which β-dominance has been noted and β-stripping applied. In that application, failed sensors are commonly modelled in the elemental filters by setting all elements in the appropriate row of the measurement matrix $\mathbf{H}_j(t_i)$ to zero. This has a tendency to make the determinant of the filter-computed residual covariance matrix associated with that filter, $|\mathbf{A}_j(t_i^-)|$, smaller than the same quantity for the fully functional measurement model. Thus, for this application, the MMAE tends to be pre-disposed to declare sensor failures even if none exist, due solely to β-dominance. This effect was noted in an MMAE used for flight control sensor failures [32].

Although β stripping is not used in this research, the concept of β-dominance is discussed frequently. This research shows that the γ design parameter in the GRMMAE derived in Chapter III causes the GRMMAE to be highly susceptible to β-dominance. This fact is derived, explained, and illustrated in Chapters III and IV.

2.4.2 Probability Lower Bounding. Observe in Equation (2.41) that, if $p_k(t_i)$ goes to zero, it will remain zero for all $t > t_i$. This causes elemental filter k to be removed from the MMAE. To prevent this from happening, it is common practice to set a lower bound ϵ on $p_k(t_i)$ (typically $0.001 < \epsilon < 0.01$). If a calculated value of $p_k(t_i)$ is less than ϵ, then $p_k(t_i)$ is set equal to ϵ, and all remaining probabilities are rescaled to maintain

$\sum_{k=1}^{K} p_k(t_i) = 1$. Probability lower bounding is implemented only in Section 4.2. Aside from that example, probability lower bounding is not used in this research.

2.4.3 Kalman Filter Tuning. As mentioned in Section 2.3, the elemental filters must be carefully tuned. Addition of too much process pseudonoise to account for modeling inadequacies can have the effect of making the likelihood quotients from all elemental filters appear similar over time. This degrades the MMAE's ability to distinguish between models and converge to the correct hypothesis. In the failure detection and isolation application, this problem manifests itself as detection delays or missed detections. However, properly tuning the elemental filters reduces false and missed alarms [8, 9, 16, 27, 30, 28]. Generally, designers use *ad hoc* methods to determine appropriate elemental filter noise covariance matrices based on physical insight into the system. Assuming elemental filter k is correct, the goal in tuning the elemental filters is to choose the noise covariance matrices such that $\mathbf{r}_k^T(t_i^-)\,\mathbf{A}_k^{-1}(t_i^-)\,\mathbf{r}_k(t_i^-)$ is on the order of the measurement dimension m, while the likelihood quotients from the other elemental filters are larger than m [46]. Kalman filter tuning is used in this research.

2.4.4 Elemental Filter Gain Modulation. Related to the previous issue of elemental filter tuning, [25] derived a method for modulating the Kalman gain in the elemental filters to increase hypothesis distinguishability. Assume that the parameter of interest affects the elemental filter dynamics model. Then, to increase the MMAE's ability to distinguish differences between the elemental filters, one wants to adjust the Kalman gain in the elemental filters to emphasize the dynamics model and de-emphasize the measurement update. This is accomplished by setting the dynamics driving noise covariance matrix $\mathbf{Q}_d(t_i)$ artificially low. However, this has the drawback of overemphasizing the dynamics model during the state estimate measurement update, degrading state estimation performance.

In this method, the Kullback Discrimination Index (KDI) (similar to a distance metric, but not a true metric) is used after each measurement update to measure the "distance" between elemental filters based on the likelihood quotients from the elemental filters. Assuming that adequate state estimation performance is being achieved, if the

elemental filters are deemed "too close" for good hypothesis distinguishability (i.e., the likelihood quotients of two or more elemental filters are of similar magnitudes), then the dynamics noise covariance matrices in the elemental filters are decreased for the next measurement update. This has the effect of increasing the emphasis on the elemental filter dynamics model during the measurement update. This increased model emphasis increases the MMAE's ability to distinguish between elemental models (hypotheses). At the next measurement update, the KDI is calculated again and the process is repeated. During this process of decreasing the dynamics pseudonoise, care must be taken to ensure that $\mathbf{Q}_d(t_i)$ is not decreased to the point where adequate state estimation performance is no longer achieved. Therefore, this method modulates the elemental filter Kalman gain at each measurement to maximize hypothesis distinguishability while ensuring state estimation performance is not degraded. Elemental filter gain modulation is not used in this research.

2.4.5 Scalar Penalty Increase. A method for speeding MMAE convergence, suggested by [31], is scalar penalty increase. In this method, $\exp\{-\frac{1}{2}\mathbf{r}_k^T(t_i^-)\,\mathbf{A}_k^{-1}(t_i^-)\,\mathbf{r}_k(t_i^-)\}$ in the Gaussian densities of Equation (2.41) is replaced by $\exp\{-\alpha\mathbf{r}_k^T(t_i^-)\,\mathbf{A}_k^{-1}(t_i^-)\,\mathbf{r}_k(t_i^-)\}$ where α is some scalar larger than 1/2. Similar to the β-stripped "densities" in Section 2.4.1, these "densities" with increased scalar penalties are not proper probability densities (i.e. do not integrate to 1). However, due to the normalization terms in the denominator of Equation (2.41), the $p_k(t_i)$ values will still sum to 1 in the MMAE.

Making α greater than 1/2 assigns higher penalties to large likelihood quotients, compared with $\alpha = 1/2$. While this speeds convergence to the correct hypothesis, it can also cause large conditional probability fluctuations during convergence. If the MMAE is used for failure detection, these fluctuations may be interpreted as momentary false alarms. Thus, the α value (which is chosen through ad hoc methods) should be chosen carefully to meet some desired performance specification. Scalar penalty increase is an alternative to GRMMAE derived in this research. It can achieve faster convergence without the susceptibility to β-dominance which plagues the GRMMAE. However, care must be taken to avoid the false alarm problem induced by scalar penalty increase. An example of scalar penalty increase is given in Section 4.5.

2.4.6 Probability Smoothing. This method is used primarily for failure detection. After a system changes operating modes, the hypothesis conditional probabilities for the elemental filters will undergo transients. These transients may cause false alarms [15, 31, 33, 34, 43, 44]. To decrease these false alarms, the hypothesis conditional probabilities can be smoothed using a moving window average. The window size is generally determined empirically. As the window size increases (i.e., more samples are used to calculate the average), the resistance to false alarms increases. However, this also causes slower response to actual failures. This research does not use probability smoothing.

2.4.7 Increased Residual Propagation. Increased residual propagation [15, 31] is similar to, yet less complex than, elemental filter gain modulation. In this method, the elemental filter state estimates are propagated a few sample times without allowing measurement updates. By doing this, the likelihood quotients from incorrect elemental filters grow because the measurement updates are not correcting the state estimates toward the measurements. However, the likelihood quotient from the correct elemental filter will remain relatively small, because its hypothesized model matches the measurements. Thus, hypothesis distinguishability is increased, leading to faster convergence. The number of samples to propagate before measurement incorporation is determined empirically. The drawback to this method is that it increases hypothesis conditional probability fluctuations, which may lead to false alarms in failure detection applications [15, 31]. Additionally, state estimation performance will decrease due to the increased time between measurement updates. This research does not use increased residual propagation.

2.4.8 Interacting Multiple Model. The MMAE assumes that the unknown parameter $\mathbf{a}$ varies slowly, if at all, compared to the states. A similar estimator exists which allows the parameter to vary under certain assumptions. This estimator is the Interacting Multiple Model (IMM) [4].

The IMM is similar to the MMAE in the sense that it uses multiple parallel Kalman filters each built around a potential parameter realization $\mathbf{a}_k$. Unlike the MMAE, however, the IMM assumes that the hypothesis conditional probability vector (composed of the probabilities associated with each elemental filter) propagates according to a Markovian

model. The IMM recursively produces updated state and covariance estimates based on the previous state and covariance estimates and the previous parameter estimate. However, IMM is suboptimal, because it only produces estimates conditioned on the current and most recent past moment in time, rather than the entire past (see [4] for an explanation).

This dissertation was inspired by the GPS carrier-phase integer ambiguity resolution application. In that application, the parameter is the GPS carrier-phase integer ambiguity vector which is constant. While the IMM allows added flexibility compared to the MMAE, it is inappropriate for that application because the parameter is truly constant. Although the GRMMAE is general enough to be applied to other applications, because it was inspired by GPS carrier-phase integer ambiguity resolution, this research does not apply the IMM.

2.5 Global Positioning System

GPS is a direct sequence spread spectrum (DSSS) satellite navigation system consisting of 24 or more medium earth orbiting satellites providing navigation information to passive receivers. The system operates using the principle of trilateration. The satellite embeds precise timing information within the transmitted signal. The receiver uses signal propagation delay to calculate the satellite distance using the formula $r = v_p(t_t - t_r)$, where r is the range from the receiver to the satellite, v_p is the velocity of propagation (usually assumed to be the speed of light in a vacuum), t_t is the transmit time, and t_r is the reception time. In addition to transmitting precise timing information, each satellite transmits ephemeris information which is used to calculate the satellite's position. From knowledge of the positions of at least four satellites and the range to each satellite, a three-dimensional user position plus clock bias solution is estimated.

2.5.1 The Global Positioning System Signal. Before examining the GPS signal, consider a general DSSS signal, represented by

$$s(t) = \sqrt{2P}cos\big[\omega_0 t + \theta_c(t) + \theta_d(t)\big] \tag{2.50}$$

where $s(t)$ is the transmitted signal, P is the average carrier power, ω_0 is the carrier angular frequency, $\theta_c\,(t)$ is the spreading code waveform, and $\theta_d\,(t)$ is the data modulation.

Table 2.3 GPS Spreading Code Characteristics [22]

Parameter	**C/A Code**	**P Code**
Data Rate	50 Hz	50 Hz
Chip Rate	1.023 Mchips/sec	10.23 Mchips/sec
Code Period	1023 Chips (1 ms)	$\approx 6 \times 10^{12}$ Chips (1 week)
Code Chip Length	293.25 m	29.325 m
Carrier Band Designation and Wavelength	L1 1575.42 MHz 0.1904 m	L1, L2 1575.42, 1227.6 MHz 0.1904, 0.2444 m respectively

Assuming antipodal binary phase shift keying (BPSK) is used for the spreading and data modulations, an equivalent form of Equation (2.50) is given by

$$\begin{aligned} s(t) &= \sqrt{2P}c(t)d(t)cos(\omega_0 t) \\ c(t) &= \pm 1 \\ d(t) &= \pm 1 \end{aligned} \tag{2.51}$$

where $d(t)$ and $c(t)$ are the data and spreading waveforms, respectively.

The GPS consists of satellites transmitting two separate DSSS signals of the form of Equation (2.51). The first signal, called the Precision code (P-code), is generally encrypted[2] and reserved for military users. The second, the Coarse/Acquistion (C/A) code, is available to civilian users. The P-code is transmitted at two carrier frequencies, L1 and L2, and C/A code is transmitted at L1. The GPS signal properties are summarized in Table 2.3.

2.6 Precise GPS Positioning

GPS positions can be calculated using either the code phase or the carrier-phase for timing. The code phase (or pseudorange[3]) measurement is used for general navigation,

[2]Encrypted P-code is called "Y-code"

[3]A GPS range measurement is termed a "pseudorange" measurement because it is the range to the satellite incorporating the effect of a clock bias.

Table 2.4 GPS Code Phase and Carrier-Phase Positioning Accuracy[18]

	Standard Code Phase GPS	**Code Phase Differential GPS**	**Carrier-Phase Differential GPS Fixed Ambiguities**
Approximate 3-D Accuracy	6-10 m	1 m [36]	0.01-0.03 m

producing position estimates with errors on the order of meters. However, due to the relatively short carrier wavelength, carrier-phase can be used to produce position estimates with errors on the order of tens of centimeters. Note, though, that carrier-phase based (or any other differential GPS based) positions are positions of one point relative to another. Table 2.4 summarizes typical 3-D accuracies achieved using stand-alone code phase, differential code phase, and differential carrier-phase positioning methods[4]. Differential (code or carrier-phase) GPS is a process of using multiple receivers and/or satellites to remove common errors. Single and double difference carrier-phase differential GPS will be described shortly.

2.7 Integer Ambiguities

To gain an understanding of carrier-phase integer ambiguities, consider Figure 2.1 which shows a GPS signal propagating from satellite k to receiver A. The range to the GPS satellite is measured by counting the number of carrier-phase cycles between the satellite and the receiver, and multiplying by the carrier wavelength. It is assumed that the receiver begins tracking the satellite signal at time t_0 and continues to track until time t. Once tracking is established at t_0, the receiver counts the carrier cycles received from that point forward in time. Additionally, the receiver is able to determine where in the partial cycle it begins to track, depicted as $\phi_{frac}(t_0)$ in Figure 2.1. However, the receiver is unable to determine the integer number of cycles N which have already propagated at time t_0. This integer number of cycles is the *carrier-phase integer ambiguity* which must be resolved to

[4]Table 2.4 assumes that the intentional degradation of Selective Availability (SA) is not present. SA was turned off in 2000 and is not expected to be used again. However, the capability to use SA again does exist, in which case standard C/A code phase position accuracies would be on the order of 100 m, while the differential (code and carrier-phase) accuracies would remain the same.

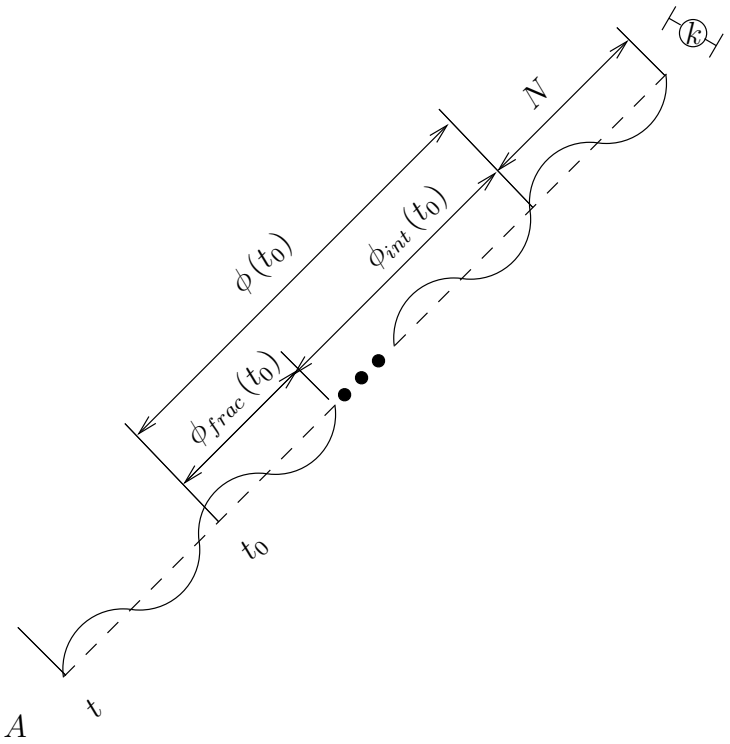


Figure 2.1 Carrier-Phase Measurement Example

determine range to the satellite. Once N is determined, it will remain constant as long as continuous carrier tracking is maintained.

The carrier-phase is calculated in receiver A by integrating the Doppler shift present in the received signal as described by Equation (2.52) [38],

$$\phi(t) = \underbrace{\int_{t_0}^{t} \Delta f_{meas}(\tau)\, d\tau + \phi(t_0)}_{\phi_{meas}(t)} + N \tag{2.52}$$

where $\phi(t)$ is the total carrier-phase cycle count at time t (corresponding to the range between the satellite and the user), $\Delta f_{meas}(\tau)$ is the received signal Doppler shift measured by the receiver at time τ, $\phi(t_0)$ is the carrier cycle count at t_0, which may include both an integer portion, $\phi_{int}(t_0)$, and a partial cycle portion, $\phi_{frac}(t_0)$, and N is the ambiguous integer carrier cycle count at the initial time. The quantity $\phi_{meas}(t) = \int_{t_0}^{t} \Delta f_{meas}(\tau)\, d\tau + \phi(t_0)$ is the carrier-phase measurement available from receiver A. One application for the GRMMAE is resolving the integer ambiguity N in Equation (2.52) and Figure 2.1.

2.8 Differential Carrier-Phase GPS

For ambiguity resolution to work efficiently, GPS range errors must be removed or reduced prior to determining the carrier-phase integer ambiguities. This is normally accomplished through differential GPS procedures known as single and double differencing. Single differencing is described first, followed by a description of double differencing. The concepts are presented for carrier-phase measurements; however, these techniques are equally valid, and similarly derived, for code phase measurements.

2.8.1 Single Difference Carrier-Phase Differential GPS. Consider two GPS receivers tracking a signal from a common GPS satellite as shown in Figure 2.2. A single difference is formed by subtracting the carrier-phase measurement of one receiver from the other. Using notation of [35], Equation (2.53) shows the carrier-phase measurement from receiver u to satellite k, in units of carrier cycles:

$$\phi_u^{(k)} = \lambda^{-1}\left[r_u^{(k)} - \left(\frac{I}{f^2}\right)_u^{(k)} + T_u^{(k)} + m_{\phi,u}^{(k)}\right] + f\left(\delta t_u - \delta t^{(k)}\right) + N_u^{(k)} + v_{\phi,u}^{(k)} \quad (2.53)$$

where

k	as a superscript, indicates the k^{th} satellite being tracked
u	as a subscript, indicates the u^{th} receiver tracking the satellite (A or B)
m_ϕ	is the carrier-phase multipath tracking error for receiver u (m)
λ	is the carrier wavelength (0.19 m for L_1 or 0.24 m for L_2)
r	is the true range from the receiver to the satellite (m)
I	is the ionospheric delay parameter, empirically determined to be 40.3 TEC ($Hz^2 \cdot$ m)[37]
T	is the tropospheric error in the measurement (m)
f	is the carrier frequency (Hz)
δt_u	is the receiver clock error (s)
$\delta t^{(k)}$	is the satellite clock error (s)

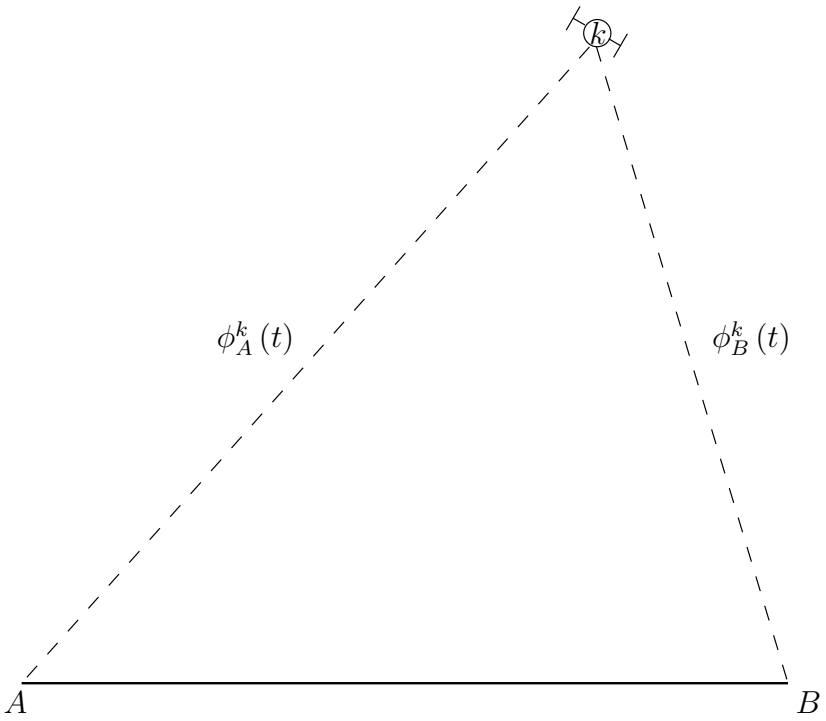


Figure 2.2 Single Difference GPS Geometry

N is the carrier-phase integer ambiguity

v_ϕ is the remaining carrier-phase measurement error (measurement noise plus possible tracking loop errors under high dynamics) (cycles)

If receivers A and B are separated by a short baseline (generally on the order of tens of kilometers), many of the errors in Equation (2.53) are highly correlated. Thus, by subtracting the phase measurement of one receiver from that of another, correlated errors can be essentially removed. Note that this difference measurement provides an indication of the relative position of the two receivers, not the range to the satellite. Thus, when differencing is used, the position of one receiver (termed a "rover") is determined relative to another receiver, which is typically at a known location (termed a "reference"). Letting $\triangle$ indicate a single difference between receivers, Equation (2.54) shows the single difference measurement between receivers A and B:

$$
\begin{aligned}
\triangle\phi_{AB}^{(k)} &= \phi_A^{(k)} - \phi_B^{(k)} \\
\triangle\phi_{AB}^{(k)} &= \left\{\lambda^{-1}\left[r_A^{(k)} - \left(\frac{I}{f^2}\right)_A^{(k)} + T_A^{(k)} + m_A^{(k)}\right] + f\left(\delta t_A - \delta t^{(k)}\right) + N_A^{(k)} + v_{\phi,A}^{(k)}\right\} \\
&\quad - \left\{\lambda^{-1}\left[r_B^{(k)} - \left(\frac{I}{f^2}\right)_B^{(k)} + T_B^{(k)} + m_B^{(k)}\right] + f\left(\delta t_B - \delta t^{(k)}\right) + N_B^{(k)} + v_{\phi,B}^{(k)}\right\} \\
\triangle\phi_{AB}^{(k)} &= \lambda^{-1}\left[\triangle r_{AB}^{(k)} - \triangle\left(\frac{I}{f^2}\right)_{AB}^{(k)} + \triangle T_{AB}^{(k)} + \triangle m_{AB}^{(k)}\right] + f\left(\delta t_A - \delta t_B\right) \\
&\quad + \triangle N_{AB}^{(k)} + \triangle v_{\phi,AB}^{(k)}
\end{aligned} \tag{2.54}
$$

In Equation (2.54), $\triangle\phi_{AB}^{(k)}$ is the single difference carrier-phase measurement between receivers A and B, $\triangle N_{AB}^{(k)}$ is the single difference phase integer ambiguity between receivers A and B, $\triangle v_{\phi,AB}^{(k)} = v_{\phi,A}^{(k)} - v_{\phi,B}^{(k)}$, and the $\triangle$ preceding the error terms indicates residual errors after differencing (i.e., the single differenced errors). Theoretically, for two receivers at the same location, these residual errors are zero[5]. As the baseline distance between the receivers increases, the error correlation between the two receivers decreases and residual errors grow. Finally, since $\triangle N_{AB}^{(k)}$ is the result of differencing two integers, it remains an integer quantity.

2.8.2 Double Difference Carrier-Phase Differential GPS. Equation (2.54) contains the receiver clock error $\delta t_A - \delta t_B$, which cannot be removed by single differencing between receivers. To obtain the most accurate relative position, this error is normally removed by forming a double difference, the difference between single difference measurements using two satellites as shown in Figure 2.3. Letting $\nabla\triangle$ represent a double difference, Equation (2.55) gives a mathematical description of double differencing between receivers A and B and satellites k and j.

$$
\nabla\triangle\phi_{AB}^{(kj)} = \triangle\phi_{AB}^{(k)} - \triangle\phi_{AB}^{(j)}
$$

[5]For this to be completely correct, the two receivers would have to use the same antenna, so that the multipath error would be identical in the two receivers.

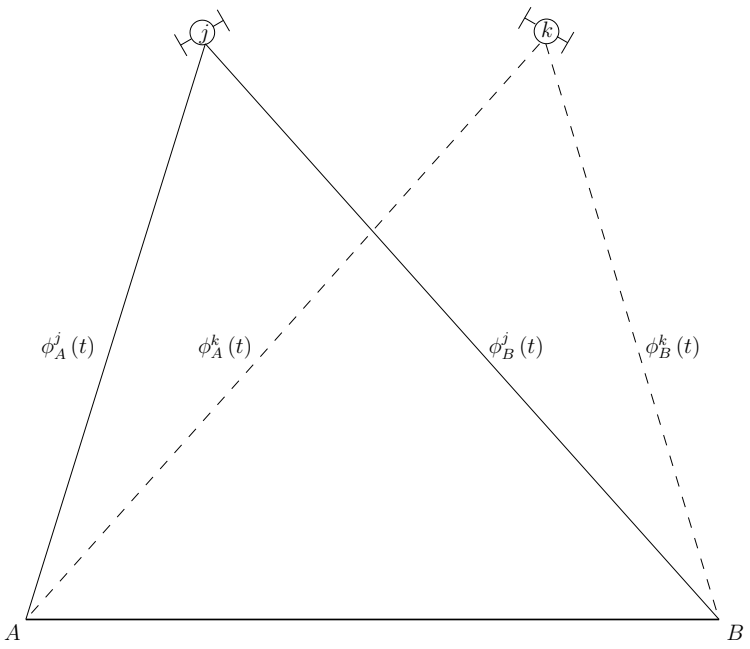


Figure 2.3 Double Difference GPS Geometry

$$
\begin{aligned}
\nabla\triangle\phi_{AB}^{(kj)} = \;& \lambda^{-1}\left[\triangle r_{AB}^{(k)} - \triangle\left(\frac{I}{f^2}\right)_{AB}^{(k)} + \triangle T_{AB}^{(k)} + \triangle m_{AB}^{(k)}\right] + f\left(\delta t_A - \delta t_B\right) + \triangle N_{AB}^{(k)} \\
+ \;& \triangle v_{\phi,AB}^{(k)} - \left\{\lambda^{-1}\left[\triangle r_{AB}^{(j)} - \triangle\left(\frac{I}{f^2}\right)_{AB}^{(j)} + \triangle T_{AB}^{(j)} + \triangle m_{AB}^{(j)}\right]\right. \\
+ \;& \left. f\left(\delta t_A - \delta t_B\right) + \triangle N_{AB}^{(j)} + \triangle v_{\phi,AB}^{(j)}\right\} \\
\nabla\triangle\phi_{AB}^{(kj)} = \;& \lambda^{-1}\left[\nabla\triangle r_{AB}^{(kj)} - \nabla\triangle\left(\frac{I}{f^2}\right)_{AB}^{(kj)} + \nabla\triangle T_{AB}^{(kj)} + \nabla\triangle m_{AB}^{(kj)}\right] \\
+ \;& \nabla\triangle N_{AB}^{(kj)} + \nabla\triangle v_{\phi,AB}^{(kj)}
\end{aligned}
\tag{2.55}
$$

In Equation (2.55), $\nabla\triangle\Phi_{AB}^{(kj)}$ is the double difference phase measurement between receivers A and B and satellites k and j. Note that, as in the single difference case, the double difference ambiguities, $\nabla\triangle N_{AB}^{(kj)}$, are again integers. Also, the $\nabla\triangle$ preceding the error terms on the right hand side of Equation (2.55) represents residual errors after double differencing. For baseline separation on the order of 10 km or less, double difference residual errors are on the order of centimeters or less [37].

2.8.3 Triple Difference Carrier-Phase Differential GPS. The previous differencing process could be extended to form triple differences. These triple differences take the difference between $\nabla\triangle\Phi_{AB}^{(kj)}$ at time t_m, and $\nabla\triangle\Phi_{AB}^{(kj)}$ at some later time t_n. Assuming continuous carrier tracking is maintained on the interval $t_m \leq t \leq t_n$, this triple difference would eliminate the integer ambiguity altogether. However, the triple difference measurement doesn't carry much information and is not generally used to obtain a differential carrier-phase positioning solution. (It is sometimes used for initialization in a static baseline case.) Because triple differencing assumes continuous carrier tracking, it can be useful for detecting breaks in carrier tracking, called "cycle slips" [35].

2.9 Modified Multiple Model Adaptive Estimator for GPS Carrier-Phase Integer Ambiguity Resolution

The technique of using multiple Kalman filters for GPS double difference carrier-phase integer ambiguity resolution was introduced early in the GPS carrier-phase ambiguity research history [21]. The concept involves building multiple extended Kalman filters, each based on a given candidate integer ambiguity vector. Then, the measurement residual vectors $\mathbf{r}_j\left(t_i^-\right)$ and residual covariance matrices $\mathbf{A}_j\left(t_i^-\right)$ from each of the elemental filters are used to determine which elemental filter is most likely to contain the correct ambiguity vector. Two researchers [18, 19] have proposed this method for integer ambiguity resolution. Their research is summarized next. Recall from Section 2.2 that using extended Kalman filters in an MMAE is an ad hoc extension which is not an optimal Bayesian estimator.

The first proposal for using multiple Kalman filters for GPS carrier-phase integer ambiguity resolution was presented in 1983 [19], prior to GPS becoming fully operational. It proposed the use of an MMAE[6] to increase the speed of the carrier-phase ambiguity resolution process in static survey applications. This research included a series of relatively simple simulations. Each case simulated double difference measurements, with a maximum of 4 GPS satellites, leading to a maximum of 3 carrier-phase ambiguities to be determined.

[6] [19] used the name "Magill Filter" after the inventor of the technique, D.T. Magill [26]

In the most demanding case, [19] demonstrated that the MMAE could be used to determine GPS carrier-phase integer ambiguities in a fairly simple, yet realistic, three dimensional GPS simulation. Since search region limiting techniques had not yet been developed, [19] relied on a brute force search of all possible integer ambiguities within 5 integers of the correct value (a total of 11 possible ambiguities per measurement), which was arbitrarily set. All simulations utilized a short baseline between receivers, and ignored many common errors including atmospheric effects and multipath. The results were encouraging, showing a simulated ability to determine the integer ambiguities with tens of minutes of measurement data, versus approximately an hour of data for batch processing techniques available at that time. However, due to the brute force search method, [19] correctly concluded that the search region in a realistic scenario would quickly grow too large. In an effort to reduce the search region, [19] derived a position and floating point ambiguity estimating Kalman filter [7] to provide an initial ambiguity estimate to the MMAE. The use of a floating point filter front end is common practice today.

This initial research demonstrated the potential to use the MMAE for GPS carrier-phase integer ambiguity resolution. However, due to computing power available at the time and the fact that GPS was in its infancy, a great deal of work remained. It was concerned only with static survey applications, and did not address the issue of kinematics. Additionally, it used only fairly simple simulations to demonstrate the technique. The technique was not demonstrated with any real GPS data.

More recently, [17, 18] proposed a modified MMAE for kinematic carrier-phase ambiguity resolution. In [17, 18] it was hypothesized that modeling inaccuracies would cause the measurement residuals to be large for all the MMAE elemental filters[8]. More specifically, recall Equation (2.41), repeated here for convenience:

$$p_k(t_i) = \frac{f_{\mathbf{z}(t_i)|\mathbf{a},\mathbf{Z}(t_{i-1})}(\mathbf{z}_i|\mathbf{a}=\mathbf{a}_k, \mathbf{Z}(t_{i-1})=\mathbf{Z}_{i-1})\, p_k(t_{i-1})}{\sum_{j=1}^{K} f_{\mathbf{z}(t_i)|\mathbf{a},\mathbf{Z}(t_{i-1})}(\mathbf{z}_i|\mathbf{a}=\mathbf{a}_j, \mathbf{Z}(t_{i-1})=\mathbf{Z}_{i-1})\, p_j(t_{i-1})} \tag{2.41}$$

[7] The floating point ambiguity Kalman filter estimates the integer ambiguity values without constraining the estimates to be integers. Therefore, the ambiguity estimates in a floating point filter are not typically integer values.

[8] This problem is related to the caution mentioned in Section 2.3 about adding dynamics pseudonoise during filter tuning.

Equation (2.41) is the conditional probability density function for the parameter $\mathbf{a}_k$, given the measurement history $\mathbf{Z}_{i-1}$. Thus, $p_k(t_i)$ is used to determine which carrier-phase integer ambiguity vector $\mathbf{a}_k$ (the ambiguity vector is the parameter in this application) is correct. The conditional densities in Equation (2.41) are given by Equation (2.47), also repeated here for convenience:

$$\begin{aligned} & f_{\mathbf{z}(t_i)|\mathbf{a},\mathbf{Z}(t_{i-1})}\left(\boldsymbol{\zeta}_i|\mathbf{a}=\mathbf{a}_k,\mathbf{Z}\left(t_{i-1}\right)=\mathbf{Z}_{i-1}\right) \\ & \quad = \frac{1}{(2\pi)^{\frac{m}{2}}\left|\mathbf{A}_k\left(t_i^-\right)\right|^{\frac{1}{2}}}\exp\left\{-\frac{1}{2}\mathbf{r}_k^T\left(t_i^-\right)\mathbf{A}_k^{-1}\left(t_i^-\right)\mathbf{r}_k\left(t_i^-\right)\right\} \end{aligned} \tag{2.47}$$

where $\mathbf{A}_k\left(t_i^-\right)=\mathbf{H}_k\left(t_i\right)\mathbf{P}_k\left(t_i^-\right)\mathbf{H}_k^T\left(t_i\right)+\mathbf{R}_k\left(t_i\right)$ is the covariance matrix of the residuals $\mathbf{r}_k\left(t_i^-\right) = \mathbf{z}_i - \mathbf{H}\left(t_i\right)\widehat{\mathbf{x}}\left(t_i^-\right)$, [9] from elemental filter k. The residual covariance $\mathbf{A}_k\left(t_i^-\right)$ is calculated from each of the K elemental filters. Equations (2.41) and (2.47) demonstrate that the MMAE relies upon the propagated state estimates $\widehat{\mathbf{x}}\left(t_i^-\right)$ to calculate measurement residuals and determine which ambiguity vector $\mathbf{a}_k$ is correct. However, the authors in [17, 18] felt the state dynamics model didn't have enough fidelity to enable the MMAE to distinguish between residuals from the elemental filters. If the dynamics model was poor, compared to the precise measurements, then the measurement residuals from all elemental filters would be large, causing MMAE failure in determining the correct ambiguity set. To correct this potential problem, [17, 18] proposed using "post-fit" residuals in the hypothesis test portion of Figure 1.2. These "post-fit" residuals are the measurement residuals calculated after measurement incorporation in each of the elemental filters. This modification appeared to fix the perceived modeling noise problem. Table 2.5 summarizes the changes made to the traditional MMAE.

The modified MMAE of Table 2.5 was built and tested using kinematic (i.e., moving) double difference carrier-phase measurements. These tests showed the modified MMAE successfully resolved GPS carrier-phase integer ambiguities. The primary contribution of this dissertation is to provide a mathematical foundation which characterizes and general-

[9] To be completely correct, [17, 18] implemented an extended Kalman filter with measurement residuals given by $\mathbf{r}_k\left(t_i^-\right) = \mathbf{z}_i - \mathbf{h}\left[\widehat{\mathbf{x}}\left(t_i^-\right), t_i\right]$. The definition above is used to be consistent with the derivation of the MMAE based on a linear Kalman filter, as presented in Section 2.3.

Table 2.5 MMAE Modification for GPS Carrier-Phase Integer Ambiguity Resolution [18]

Traditional MMAE	**Modified MMAE**
$\mathbf{r}_k\left(t_i^-\right) = \mathbf{z}\left(t_i\right) - \mathbf{h}\left[\widehat{\mathbf{x}}_k\left(t_i^-\right), t_i\right]$	$\mathbf{r}_k\left(t_i^+\right) = \mathbf{z}\left(t_i\right) - \mathbf{h}\left[\widehat{\mathbf{x}}_k\left(t_i^+\right), t_i\right]$
$\mathbf{A}_k\left(t_i^-\right) = \mathbf{H}_k\left(t_i\right)\mathbf{P}_k\left(t_i^-\right)\mathbf{H}_k^T\left(t_i\right) + \mathbf{R}\left(t_i\right)$	$\mathbf{A}_k\left(t_i^+\right) = \mathbf{H}_k\left(t_i\right)\mathbf{P}_k\left(t_i^+\right)\mathbf{H}_k^T\left(t_i\right) + \mathbf{R}\left(t_i\right)$

izes the modified MMAE. While the modified MMAE was designed for GPS carrier-phase integer ambiguity resolution, development of the GRMMAE is not restricted to any particular application. The modified MMAE of [17, 18] is one specific example of the GRMMAE developed in Chapter III.

2.10 Chapter Summary

The background necessary to understand the GRMMAE derivation in Chapter III was provided in this chapter. The Kalman filter was derived as the basic building block for the Multiple Model Adaptive Estimator. The extended Kalman filter was developed as it is frequently used in an ad hoc extension of the MMAE. The hypothesis conditional probability formula for the MMAE was derived and several MMAE performance enhancements were presented. Since this dissertation is inspired by the use of a modified MMAE for GPS carrier-phase integer ambiguity resolution, an introduction to GPS was presented. Carrier-phase integer ambiguities were described and single and double difference carrier-phase differential GPS were discussed. The chapter concluded by describing the modified MMAE for GPS carrier-phase integer ambiguity resolution. The contribution of this dissertation is to provide a sound mathematical formulation which generalizes this modified MMAE.

III. Theory

In this chapter, the Generalized Residual Multiple Model Adaptive Estimator is derived. A "generalized residual," which is a linear combination of traditional and post-fit residuals, is defined. Properties of the generalized residuals from elemental filters based on both the correct and incorrect hypotheses are given for some common MMAE applications. The generalized residuals and their properties are then used to re-derive an MMAE based on the generalized residual, rather than the traditional residual. The newly derived MMAE is called the Generalized Residual Multiple Model Adaptive Estimator (GRMMAE). The GRMMAE has exactly the same form and purpose as the traditional MMAE. The only functional difference between the two is the formula used for hypothesis conditional probability calculation. This derivation assumes the GRMMAE elemental filters are linear Kalman filters. The effect of using an extended Kalman filter for the GPS carrier phase integer ambiguity resolution application is discussed in Chapter IV.

Each section in this chapter provides details of a specific research contribution resulting from this dissertation. The entirety of this chapter is the primary contribution of providing a mathematical foundation and generalization for the modified MMAE. In addition to its application to the modified MMAE, the derivation provided is general enough to apply to a traditional MMAE or any other MMAE using a generalized residual.

3.1 Generalized Residual

Begin by defining the generalized residual vector as a linear combination of traditional and post-fit residuals

$$\mathbf{r}^*(t_i) = \gamma \mathbf{r}\left(t_i^-\right) + (1 - \gamma)\,\mathbf{r}\left(t_i^+\right) \tag{3.1}$$

where $\mathbf{r}^*(t_i)$ is the generalized residual, $\mathbf{r}\left(t_i^-\right)$ is the traditional Kalman filter residual, $\mathbf{r}\left(t_i^+\right)$ is the post-fit residual, both defined in Table 2.5, and γ is a real, scalar, user-defined GRMMAE design parameter. Although γ is not constrained to be in the region from 0 to 1,

inclusively, $\gamma \in [0,1]$ has the most physical meaning. It is anticipated that most engineering applications will use $\gamma \in [0,1]$. Finally, throughout this dissertation, a superscript $*$ denotes variables based on the generalized residual. The $*$ notation does not denote the complex conjugate anywhere in this dissertation.

A simple derivation yields a matrix transformation from the traditional to the generalized residual

$$\begin{aligned}
\mathbf{r}^*(t_i) &= \mathbf{z}(t_i) - \gamma \mathbf{H}(t_i)\,\widehat{\mathbf{x}}\left(t_i^-\right) - (1-\gamma)\,\mathbf{H}(t_i)\,\widehat{\mathbf{x}}\left(t_i^+\right) \\
&= \mathbf{z}(t_i) - \gamma \mathbf{H}(t_i)\,\widehat{\mathbf{x}}\left(t_i^-\right) \\
&\quad - (1-\gamma)\,\mathbf{H}(t_i)\left[\widehat{\mathbf{x}}\left(t_i^-\right) + \mathbf{K}(t_i)\left[\mathbf{z}(t_i) - \mathbf{H}(t_i)\,\widehat{\mathbf{x}}\left(t_i^-\right)\right]\right] \\
&= \left[\mathbf{I} - (1-\gamma)\,\mathbf{H}(t_i)\,\mathbf{K}(t_i)\right]\left[\mathbf{z}(t_i) - \mathbf{H}(t_i)\,\widehat{\mathbf{x}}\left(t_i^-\right)\right] \\
\mathbf{r}^*(t_i) &= \mathbf{T}(t_i)\,\mathbf{r}\left(t_i^-\right) \qquad (3.2)
\end{aligned}$$

where $\mathbf{T}(t_i) = \left[\mathbf{I} - (1-\gamma)\,\mathbf{H}(t_i)\,\mathbf{K}(t_i)\right]$ is the transformation that interrelates $\mathbf{r}^*(t_i)$ and $\mathbf{r}\left(t_i^-\right)$. The transformation $\mathbf{T}(t_i)$ and design parameter γ are discussed further in Sections 3.4 and 3.7, respectively.

3.2 Generalized Residual Properties for the Correct Elemental Filter

Generalized residual vector properties are derived in this section, assuming that the k^{th} GRMMAE elemental filter contains the correct hypothesis. From Equation (3.2),

$$\mathbf{r}_k^*(t_i) = \mathbf{T}_k(t_i)\,\mathbf{r}_k\left(t_i^-\right)$$

where the subscript k indicates variables associated with the k^{th} elemental filter. Since the k^{th} filter is assumed correct, $\mathbf{r}_k\left(t_i^-\right)$ is Gaussian [29]. The generalized residual $\mathbf{r}_k^*(t_i)$ is a linear function of $\mathbf{r}_k\left(t_i^-\right)$, and therefore, also Gaussian. Thus, $\mathbf{r}_k^*(t_i)$ is completely defined by its mean and covariance. Given Equation (3.2), the conditional mean of the generalized residual from the correct elemental filter is

$$
\begin{aligned}
\mathbf{m}_{\mathbf{r}_k^*(t_i)} &= E\left[\mathbf{r}^*\left(t_i\right) \mid \mathbf{a} = \mathbf{a}_k,\ \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\right] \\
&= E\Big[\mathbf{T}\left(t_i\right)\mathbf{r}\left(t_i^-\right) \mid \mathbf{a} = \mathbf{a}_k,\ \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\Big] \\
&= \mathbf{T}_k\left(t_i\right) E\Big[\mathbf{r}\left(t_i^-\right) \mid \mathbf{a} = \mathbf{a}_k,\ \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\Big] \\
\mathbf{m}_{\mathbf{r}_k^*(t_i)} &= \mathbf{0}
\end{aligned} \tag{3.3}
$$

where the final line above comes from $E\left[\mathbf{r}\left(t_i^-\right) \mid \mathbf{a} = \mathbf{a}_k,\ \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\right] = \mathbf{0}$ [29]. Notice that, similar to the development in Section 2.3, the measurement history at t_{i-1} is used in Equation (3.3).

Since $\mathbf{r}_k^*\left(t_i\right)$ is zero-mean, the conditional covariance is given by

$$
\begin{aligned}
\mathbf{A}_k^*\left(t_i\right) &= E\Big[\mathbf{r}^*\left(t_i\right)\mathbf{r}^{*^T}\left(t_i\right) \mid \mathbf{a} = \mathbf{a}_k,\ \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\Big] \\
&= E\Big[\mathbf{T}\left(t_i\right)\mathbf{r}\left(t_i^-\right)\left(\mathbf{T}\left(t_i\right)\mathbf{r}\left(t_i^-\right)\right)^T \mid \mathbf{a} = \mathbf{a}_k,\ \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\Big] \\
&= \mathbf{T}_k\left(t_i\right) E\Big[\mathbf{r}\left(t_i^-\right)\mathbf{r}^T\left(t_i^-\right) \mid \mathbf{a} = \mathbf{a}_k,\ \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\Big]\mathbf{T}_k^T\left(t_i\right) \\
&= \mathbf{T}_k\left(t_i\right)\mathbf{A}_k\left(t_i^-\right)\mathbf{T}_k^T\left(t_i\right) \\
\mathbf{A}_k^*\left(t_i\right) &= \left[\mathbf{I} - (1-\gamma)\,\mathbf{H}_k\left(t_i\right)\mathbf{K}_k\left(t_i\right)\right]\left[\mathbf{H}_k\left(t_i\right)\mathbf{P}_k\left(t_i^-\right)\mathbf{H}_k^T\left(t_i\right) + \mathbf{R}\left(t_i\right)\right] \\
&\quad \left[\mathbf{I} - (1-\gamma)\,\mathbf{H}_k\left(t_i\right)\mathbf{K}_k\left(t_i\right)\right]^T
\end{aligned} \tag{3.4}
$$

where the substitution

$$
E\Big[\mathbf{r}\left(t_i^-\right)\mathbf{r}^T\left(t_i^-\right) \mid \mathbf{a} = \mathbf{a}_k,\ \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\Big] = \left[\mathbf{H}_k\left(t_i\right)\mathbf{P}_k\left(t_i^-\right)\mathbf{H}_k^T\left(t_i\right) + \mathbf{R}\left(t_i\right)\right]
$$

is given in [30], and $\mathbf{T}_k\left(t_i\right)$ is brought outside the expectation because it is not a random matrix.

Consider the following specific cases for Equation (3.4)

$$\mathbf{A}_k^*(t_i) \;=$$

$$\begin{cases} \mathbf{H}_k(t_i)\,\mathbf{P}_k\left(t_i^-\right)\mathbf{H}_k^T(t_i) \;+\; \mathbf{R}(t_i) & \gamma = 1 \\ \left[\mathbf{I} - \mathbf{H}_k(t_i)\,\mathbf{K}_k(t_i)\right]\left[\mathbf{H}_k(t_i)\,\mathbf{P}_k\left(t_i^-\right)\mathbf{H}_k^T(t_i) + \mathbf{R}(t_i)\right]\left[\mathbf{I} - \mathbf{H}_k(t_i)\,\mathbf{K}_k(t_i)\right]^T & \gamma = 0 \end{cases} \tag{3.5}$$

The $\gamma = 1$ case is the traditional MMAE and the result of Equation (3.5) exactly matches the traditional filter-computed residual covariance [30]

$$\mathbf{A}_k(t_i) \;=\; \left[\mathbf{H}_k(t_i)\,\mathbf{P}_k\left(t_i^-\right)\mathbf{H}^T(t_i) \;+\; \mathbf{R}(t_i)\right]$$

The $\gamma = 0$ case is a GRMMAE using post-fit residuals. This result is easily confirmed using Equation (3.2).

3.3 Generalized Residual Properties for Incorrect Elemental Filters

Generalized residual properties for the correct GRMMAE elemental filter were derived in the previous section. A similarly general derivation cannot be accomplished for the incorrect elemental filters. However, for specific applications, generalized residual properties from the incorrect elemental filters can be derived. This is accomplished for four different types of MMAE problems: 1) a measurement bias problem; 2) an unknown equivalent discrete-time dynamics noise covariance problem; 3) an unknown dynamics matrix problem; 4) an unknown measurement noise covariance problem. The procedures shown in these examples can be applied in a similar manner to other types of problems as well.

3.3.1 Measurement Bias Parameter. When the parameter to be estimated is a constant measurement bias vector, the residual properties for both the correct and incorrect elemental filters can be derived. An example of such a parameter is the GPS carrier-phase integer ambiguity vector (the problem originally motivating this research). Table 3.1 summarizes the relationship between the elemental filters for the key matrices and vectors in the measurement bias scenario under the previous assumption that the elemental

Table 3.1 Relationship Among Elemental Filter Matrices and Vectors for Unknown Measurement Bias Application

Same in All Elemental Filters	**Different in Each Elemental Filter**
$\mathbf{z}(t_i)$	$\tilde{\mathbf{z}}(t_i)$
$\mathbf{B}_d(t_i)$	$\widehat{\mathbf{x}}(t_i^-)$ (after first measurement)
$\mathbf{u}(t_i)$	$\widehat{\mathbf{x}}(t_i^+)$
$\mathbf{Q}_d(t_i)$	$\mathbf{r}^*(t_i)$
$\mathbf{H}(t_i)$	
$\mathbf{R}(t_i)$	
$\mathbf{K}(t_i)$	
$\mathbf{P}(t_i^-)$	
$\mathbf{P}(t_i^+)$	
$\mathbf{A}^*(t_i)$	
$\mathbf{T}(t_i)$	
$\mathbf{\Phi}(t_i, t_{i-1})$	

filters are based on linear system models. Chapter IV discusses the impact of using a nonlinear measurement model for GPS carrier-phase integer ambiguity resolution on the (linear model) derivation presented here.

Assume the correct parameter is contained in the k^{th} elemental filter. The sensor measurement can be modeled by

$$\mathbf{z}(t_i) \;= \mathbf{z}_d(t_i) \;+\; \mathbf{n}_k \tag{3.6}$$

where $\mathbf{z}(t_i)$ is the measurement output from the sensor, $\mathbf{z}_d(t_i)$ is the desired (unbiased) measurement, and $\mathbf{n}_k$ is the unknown, constant measurement bias vector (i.e., the parameter to be determined). Note that the desired measurement $\mathbf{z}_d(t_i)$ is a noise-corrupted measurement. It is termed "desired" here because it is unbiased.

For ease of implementation, each elemental filter can be designed to operate on a corrected measurement given by [18, 20]

$$\begin{aligned} \tilde{\mathbf{z}}_j(t_i) \;&= \mathbf{z}(t_i) \;-\; \mathbf{n}_j \\ \tilde{\mathbf{z}}_j(t_i) \;&= \mathbf{z}_d(t_i) \;+\; \mathbf{n}_k - \;\mathbf{n}_j \end{aligned} \tag{3.7}$$

where $\mathbf{n}_j$ is the constant measurement bias hypothesized in the j^{th} elemental filter and $\mathbf{n}_k$ is assumed to be the correct bias.

Using the corrected measurement of Equation (3.7) leads to all elemental filters having the same structure but operating on different measurements. It is particularly important to note that

$$\begin{aligned} \mathbf{H}_k(t_i) &= \mathbf{H}_j(t_i) = \mathbf{H}(t_i) \\ &\text{and} \\ \mathbf{K}_k(t_i) &= \mathbf{K}_j(t_i) = \mathbf{K}(t_i) \end{aligned} \tag{3.8}$$

Equation (3.8) is valid because of the way the elemental filters were constructed for this application. It implies that the transformation matrix $\mathbf{T}(t_i)$ is the same in every elemental filter.

Assuming the measurement history up to and including t_{i-1} is available, the residual vector for any elemental filter, given Equation (3.7), is

$$\begin{aligned} \mathbf{r}_j^*(t_i) &= \tilde{\mathbf{z}}_j(t_i) - \gamma\mathbf{H}(t_i)\widehat{\mathbf{x}}_j\left(t_i^-\right) - (1-\gamma)\mathbf{H}(t_i)\widehat{\mathbf{x}}_j\left(t_i^+\right) \\ &= \mathbf{z}_d(t_i) + \mathbf{n}_k - \mathbf{n}_j - \gamma\mathbf{H}(t_i)\widehat{\mathbf{x}}_j\left(t_i^-\right) \\ &\quad - (1-\gamma)\mathbf{H}(t_i)\left\{\widehat{\mathbf{x}}_j\left(t_i^-\right) + \mathbf{K}(t_i)\left[\mathbf{z}_d(t_i) + \mathbf{n}_k - \mathbf{n}_j - \mathbf{H}(t_i)\widehat{\mathbf{x}}_j\left(t_i^-\right)\right]\right\} \\ \mathbf{r}_j^*(t_i) &= \mathbf{z}_d(t_i) + \boldsymbol{\Delta}\mathbf{n}_{kj} - \gamma\mathbf{H}(t_i)\widehat{\mathbf{x}}_j\left(t_i^-\right) \\ &\quad - (1-\gamma)\mathbf{H}(t_i)\left\{\widehat{\mathbf{x}}_j\left(t_i^-\right) + \mathbf{K}(t_i)\left[\mathbf{z}_d(t_i) + \boldsymbol{\Delta}\mathbf{n}_{kj} - \mathbf{H}(t_i)\widehat{\mathbf{x}}_j\left(t_i^-\right)\right]\right\} \end{aligned} \tag{3.9}$$

where $\boldsymbol{\Delta}\mathbf{n}_{kj} = \mathbf{n}_k - \mathbf{n}_j$ is the difference between the hypothesized measurement bias in elemental filter j and the correct measurement bias in elemental filter k. Equation (3.9) can be simplified to

$$
\begin{aligned}
\mathbf{r}_j^*(t_i) &= \left[\mathbf{I} - (1-\gamma)\,\mathbf{H}(t_i)\,\mathbf{K}(t_i)\right]\left[\mathbf{z}_d(t_i) - \mathbf{H}(t_i)\,\widehat{\mathbf{x}}_j\left(t_i^-\right) + \boldsymbol{\Delta}\mathbf{n}_{kj}\right] \\
&= \mathbf{T}(t_i)\left[\mathbf{z}_d(t_i) - \mathbf{H}(t_i)\,\widehat{\mathbf{x}}_k\left(t_i^-\right) + \mathbf{H}(t_i)\,\widehat{\mathbf{x}}_k\left(t_i^-\right) - \mathbf{H}(t_i)\,\widehat{\mathbf{x}}_j\left(t_i^-\right) + \boldsymbol{\Delta}\mathbf{n}_{kj}\right] \\
&= \mathbf{T}(t_i)\left[\mathbf{r}_k\left(t_i^-\right) + \mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right) - \widehat{\mathbf{x}}_j\left(t_i^-\right)\right) + \boldsymbol{\Delta}\mathbf{n}_{kj}\right] \\
\mathbf{r}_j^*(t_i) &= \begin{cases} \mathbf{T}(t_i)\,\mathbf{r}_k\left(t_i^-\right) & j = k \\ \mathbf{T}(t_i)\left[\mathbf{r}_k\left(t_i^-\right) + \mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right) - \widehat{\mathbf{x}}_j\left(t_i^-\right)\right) + \boldsymbol{\Delta}\mathbf{n}_{kj}\right] & \forall j \end{cases}
\end{aligned}
\tag{3.10}
$$

In the third line of Equation (3.10), $\mathbf{r}_k\left(t_i^-\right)$ is equated to $\mathbf{z}_d(t_i) - \mathbf{H}(t_i)\,\widehat{\mathbf{x}}_k\left(t_i^-\right)$. That relationship only holds for the k^{th} (i.e., correct) elemental filter. In general, $\mathbf{r}_j\left(t_i^-\right) = \tilde{\mathbf{z}}_j(t_i) - \mathbf{H}(t_i)\,\widehat{\mathbf{x}}_j\left(t_i^-\right)$. In the special case when $j = k$, $\tilde{\mathbf{z}}_j(t_i) = \mathbf{z}_d(t_i)$, as seen in Equation 3.7, which leads to $\mathbf{r}_k\left(t_i^-\right) = \mathbf{z}_d(t_i) - \mathbf{H}(t_i)\,\widehat{\mathbf{x}}_k\left(t_i^-\right)$.

Given $E\left[\mathbf{r}_k\left(t_i^-\right) \mid \mathbf{a} = \mathbf{a}_k,\ \mathbf{Z}(t_{i-1}) = \mathbf{Z}_{i-1}\right] = \mathbf{0}$ [29] and Equation (3.10), the expected value of $\mathbf{r}_j^*(t_i)$ is

$$
\begin{aligned}
\mathbf{m}_{\mathbf{r}_j^*}(t_i) &= E\left[\mathbf{r}_j^*(t_i) \mid \mathbf{a} = \mathbf{a}_k,\ \mathbf{Z}(t_{i-1}) = \mathbf{Z}_{i-1}\right] \\
\mathbf{m}_{\mathbf{r}_j^*}(t_i) &= \begin{cases} \mathbf{0} & j = k \\ \mathbf{T}(t_i)\left[\mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right) - \widehat{\mathbf{x}}_j\left(t_i^-\right)\right) + \boldsymbol{\Delta}\mathbf{n}_{kj}\right] & \forall j \end{cases}
\end{aligned}
\tag{3.11}
$$

The covariance of $\mathbf{r}_j^*(t_i)$ is

$$
\begin{aligned}
\mathbf{A}_j^*(t_i) &= E\left[\mathbf{r}_j^*(t_i)\,\mathbf{r}_j^{*^T}(t_i) \mid \mathbf{a} = \mathbf{a}_k,\ \mathbf{Z}(t_{i-1}) = \mathbf{Z}_{i-1}\right] - \mathbf{m}_{\mathbf{r}_j^*}(t_i)\,\mathbf{m}_{\mathbf{r}_j^*}^T(t_i) \\
&= \mathbf{T}(t_i)\,E\Big\{\left[\mathbf{r}_k\left(t_i^-\right) + \mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right) - \widehat{\mathbf{x}}_j\left(t_i^-\right)\right) + \boldsymbol{\Delta}\mathbf{n}_{kj}\right] \\
&\quad \left[\mathbf{r}_k\left(t_i^-\right) + \mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right) - \widehat{\mathbf{x}}_j\left(t_i^-\right)\right) + \boldsymbol{\Delta}\mathbf{n}_{kj}\right]^T \mid \mathbf{a} = \mathbf{a}_k,\ \mathbf{Z}(t_{i-1}) = \mathbf{Z}_{i-1}\Big\}\mathbf{T}^T(t_i) \\
&\quad - \mathbf{m}_{\mathbf{r}_j^*}(t_i)\,\mathbf{m}_{\mathbf{r}_j^*}^T(t_i)
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{A}_j^*(t_i) \;=\; & \mathbf{T}(t_i)\Big\{E\Big[\mathbf{r}_k\left(t_i^-\right)\mathbf{r}_k^T\left(t_i^-\right) \mid \mathbf{a} \;=\; \mathbf{a}_k,\; \mathbf{Z}(t_{i-1}) \;=\; \mathbf{Z}_{i-1}\Big] \\
& +\; E\Big[\mathbf{r}_k\left(t_i^-\right) \mid \mathbf{a} \;=\; \mathbf{a}_k,\; \mathbf{Z}(t_{i-1}) \;=\; \mathbf{Z}_{i-1}\Big]\left(\widehat{\mathbf{x}}_k\left(t_i^-\right)-\widehat{\mathbf{x}}_j\left(t_i^-\right)\right)^T\mathbf{H}(t_i)^T \\
& +\; E\Big[\mathbf{r}_k\left(t_i^-\right) \mid \mathbf{a} \;=\; \mathbf{a}_k,\; \mathbf{Z}(t_{i-1}) \;=\; \mathbf{Z}_{i-1}\Big]\boldsymbol{\Delta}\mathbf{n}_{kj}^T \\
& +\; \mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right)-\widehat{\mathbf{x}}_j\left(t_i^-\right)\right)E\Big[\mathbf{r}_k^T\left(t_i^-\right) \mid \mathbf{a} \;=\; \mathbf{a}_k,\; \mathbf{Z}(t_{i-1}) \;=\; \mathbf{Z}_{i-1}\Big] \\
& +\; \mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right)-\widehat{\mathbf{x}}_j\left(t_i^-\right)\right)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right)-\widehat{\mathbf{x}}_j\left(t_i^-\right)\right)^T\mathbf{H}(t_i)^T \\
& +\; \mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right)-\widehat{\mathbf{x}}_j\left(t_i^-\right)\right)\boldsymbol{\Delta}\mathbf{n}_{kj}^T \\
& +\; \boldsymbol{\Delta}\mathbf{n}_{kj}E\Big[\mathbf{r}_k^T\left(t_i^-\right) \mid \mathbf{a} \;=\; \mathbf{a}_k,\; \mathbf{Z}(t_{i-1}) \;=\; \mathbf{Z}_{i-1}\Big] \\
& +\; \boldsymbol{\Delta}\mathbf{n}_{kj}\left(\widehat{\mathbf{x}}_k\left(t_i^-\right)-\widehat{\mathbf{x}}_j\left(t_i^-\right)\right)^T\mathbf{H}(t_i)^T \;+\; \boldsymbol{\Delta}\mathbf{n}_{kj}\boldsymbol{\Delta}\mathbf{n}_{kj}^T\Big\}\mathbf{T}(t_i) \;-\; \mathbf{m}_{\mathbf{r}_j^*}(t_i)\,\mathbf{m}_{\mathbf{r}_j^*}^T(t_i)
\end{aligned}
\tag{3.12}
$$

The mean and covariance of the traditional residual from the correct elemental filter $\mathbf{r}_k\left(t_i^-\right)$ are used to simplify Equation (3.12):

$$
\begin{aligned}
\mathbf{A}_j^*(t_i) \;=\; & \mathbf{T}(t_i)\Big\{\mathbf{A}_k\left(t_i^-\right) \;+\; \left[\mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right)-\widehat{\mathbf{x}}_j\left(t_i^-\right)\right) \;+\; \boldsymbol{\Delta}\mathbf{n}_{kj}\right] \\
& \left[\mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right)-\widehat{\mathbf{x}}_j\left(t_i^-\right)\right) \;+\; \boldsymbol{\Delta}\mathbf{n}_{kj}\right]^T\Big\}\mathbf{T}^T(t_i) \;-\; \mathbf{m}_{\mathbf{r}_j^*}\mathbf{m}_{\mathbf{r}_j^*}^T \\
=\; & \mathbf{A}_k^*(t_i) \;+\; \mathbf{T}(t_i)\left[\mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right)-\widehat{\mathbf{x}}_j\left(t_i^-\right)\right) \;+\; \boldsymbol{\Delta}\mathbf{n}_{kj}\right] \\
& \left[\mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right)-\widehat{\mathbf{x}}_j\left(t_i^-\right)\right) \;+\; \boldsymbol{\Delta}\mathbf{n}_{kj}\right]^T\mathbf{T}^T(t_i) \\
& -\; \mathbf{T}(t_i)\left[\mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right)-\widehat{\mathbf{x}}_j\left(t_i^-\right)\right) \;+\; \boldsymbol{\Delta}\mathbf{n}_{kj}\right] \\
& \left[\mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right)-\widehat{\mathbf{x}}_j\left(t_i^-\right)\right) \;+\; \boldsymbol{\Delta}\mathbf{n}_{kj}\right]^T\mathbf{T}^T(t_i) \\
\mathbf{A}_j^*(t_i) \;=\; & \mathbf{A}_k^*(t_i) \;=\; \mathbf{T}(t_i)\,\mathbf{A}_k\left(t_i^-\right)\mathbf{T}^T(t_i)
\end{aligned}
\tag{3.13}
$$

Thus, the generalized residuals from all elemental filters will have the same filter-computed covariance and differing means when the parameter of interest is a measurement bias.

3.3.2 Residual Properties for Varying $\mathbf{Q}_d(t_i)$ Parameter. Similar to the previous subsection, residual properties can also be established in the more traditional MMAE application of determining the correct equivalent discrete-time dynamics driving noise co-

Table 3.2 Relationship Among Elemental Filter Matrices and Vectors for Unknown $\mathbf{Q}_d$ Application

Same in All Elemental Filters	**Different in Each Elemental Filter**
$\mathbf{\Phi}(t_i, t_{i-1})$	$\widehat{\mathbf{x}}(t_i^-)$
$\mathbf{B}_d(t_i)$	$\widehat{\mathbf{x}}(t_i^+)$
$\mathbf{u}(t_i)$	$\mathbf{P}(t_i^-)$
$\mathbf{H}(t_i)$	$\mathbf{P}(t_i^+)$
$\mathbf{R}(t_i)$	$\mathbf{K}(t_i)$
$\mathbf{z}(t_i)$	$\mathbf{r}^*(t_i)$
	$\mathbf{A}^*(t_i)$
	$\mathbf{T}(t_i)$
	$\mathbf{Q}_d(t_i)$

variance $\mathbf{Q}_d(t_i)$. In this problem, each elemental filter is built with a different hypothesized $\mathbf{Q}_d(t_i)$ matrix. Table 3.2 summarizes the relationship among the key elemental filter matrices and vectors for this scenario, assuming that the elemental filter models are linear.

Assume elemental filter k is based on the correct hypothesis. To calculate the mean of $\mathbf{r}_j^*(t_i)$ first define $\mathbf{r}_j^*(t_i)$ in terms of $\mathbf{r}_k(t_i^-)$

$$\begin{aligned}
\mathbf{r}_j^*(t_i) &= \left[\mathbf{I} - (1 - \gamma)\mathbf{H}_j(t_i)\mathbf{K}_j(t_i)\right]\mathbf{r}_j(t_i^-) \\
&= \mathbf{T}_j(t_i)\mathbf{r}_j(t_i^-) \\
\mathbf{r}_j^*(t_i) &= \mathbf{T}_j(t_i)\left[\mathbf{r}_k(t_i^-) + \mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_k(t_i^-) - \widehat{\mathbf{x}}_j(t_i^-)\right)\right] \qquad (3.14)
\end{aligned}$$

Given Equation (3.14), the mean and covariance of $\mathbf{r}_j^*(t_i)$ are calculated from the mean and covariance of $\mathbf{r}_k(t_i^-)$.

$$\begin{aligned}
\mathbf{m}_{\mathbf{r}_j^*}(t_i) &= E\left[\mathbf{r}_j^*(t_i) \mid \mathbf{a} = \mathbf{a}_k,\ \mathbf{Z}(t_{i-1}) = \mathbf{Z}_{i-1}\right] \\
&= E\left[\mathbf{T}_j(t_i)\left[\mathbf{r}_k(t_i^-) + \mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_k(t_i^-) - \widehat{\mathbf{x}}_j(t_i^-)\right)\right] \mid \mathbf{a} = \mathbf{a}_k,\ \mathbf{Z}(t_{i-1}) = \mathbf{Z}_{i-1}\right] \\
\mathbf{m}_{\mathbf{r}_j^*}(t_i) &= \mathbf{T}_j(t_i)\mathbf{H}(t_i)\left[\widehat{\mathbf{x}}_k(t_i^-) - \widehat{\mathbf{x}}_j(t_i^-)\right]
\end{aligned}$$

$$\begin{aligned}
\mathbf{A}_j^*(t_i) &= E\left[\mathbf{r}_j^*(t_i)\mathbf{r}_j^{*T}(t_i) \mid \mathbf{a} = \mathbf{a}_k,\ \mathbf{Z}(t_{i-1}) = \mathbf{Z}_{i-1}\right] - \mathbf{m}_{\mathbf{r}_j^*}(t_i)\mathbf{m}_{\mathbf{r}_j^*}^T(t_i) \\
\mathbf{A}_j^*(t_i) &= \mathbf{T}_j(t_i)\mathbf{A}_k(t_i^-)\mathbf{T}_j^T(t_i) \qquad (3.15)
\end{aligned}$$

As in the measurement bias application, the generalized residual vectors will have differing means, dependent upon the elemental filter. However, unlike the measurement bias application, the filter-computed residual covariance will also vary among the elemental filters because the transformation matrix $\mathbf{T}_j(t_i)$ varies among the elemental filters.

3.3.3 Residual Properties for Varying $\mathbf{F}(t)$ Parameter. The linear difference equation for system dynamics in Equation (2.1) is calculated from the differential equation [29]

$$\dot{\mathbf{x}}(t) = \mathbf{F}(t)\,\mathbf{x}(t) + \mathbf{B}(t)\,\mathbf{u}(t) + \mathbf{G}(t)\,\mathbf{w}(t) \tag{3.16}$$

where $\mathbf{F}(t)$ is the system dynamics matrix. The state transition matrix, $\mathbf{\Phi}(t_i, t_{i-1})$, in Equation (2.1) is the matrix which satisfies the differential equation and initial condition:

$$\begin{aligned} \frac{d\left[\mathbf{\Phi}(t, t_0)\right]}{dt} &= \mathbf{F}(t)\,\mathbf{\Phi}(t, t_0) \\ \mathbf{\Phi}(t_0, t_0) &= \mathbf{I} \end{aligned} \tag{3.17}$$

For a time invariant system (i.e., constant $\mathbf{F}(t)$), the state transition matrix can be calculated using the matrix exponential

$$\mathbf{\Phi}(t, t_0) = \mathbf{\Phi}(t - t_0) = e^{\mathbf{F}(t-t_0)} \tag{3.18}$$

Thus, the system dynamics matrix $\mathbf{F}(t)$ is used to define the linear difference equation of Equation (2.1). With this in mind, an example of a GRMMAE with an unknown $\mathbf{F}(t)$ is presented next.

In this scenario, it is assumed that the matrix describing the system dynamics, $\mathbf{F}(t)$ contains an unknown parameter. Such a case arises when the $\mathbf{F}(t)$ matrix contains a time constant, and this approach useful for target identification. Different-valued time constants are hypothesized for various targets, with an observed target being classified by

Table 3.3 Relationship Among Elemental Filter Matrices and Vectors for Unknown $\mathbf{F}(t_i)$ Application

Same in All Elemental Filters	**Different in Each Elemental Filter**
$\mathbf{u}(t_i)$	$\mathbf{B}_d(t_i)$
$\mathbf{H}(t_i)$	$\widehat{\mathbf{x}}(t_i^-)$
$\mathbf{R}(t_i)$	$\mathbf{\Phi}(t_i, t_{i-1})$
$\mathbf{z}(t_i)$	$\mathbf{P}(t_i^-)$
	$\mathbf{Q}_d(t_i)$
	$\mathbf{K}(t_i)$
	$\widehat{\mathbf{x}}(t_i^+)$
	$\mathbf{P}(t_i^+)$
	$\mathbf{T}(t_i)$
	$\mathbf{r}^*(t_i)$

determining the most probable time constant. Since the state transition matrix $\mathbf{\Phi}(t_i, t_{i-1})$, is calculated directly from $\mathbf{F}(t)$, this scenario is equivalent to an unknown $\mathbf{\Phi}(t_i, \tau)$ scenario. Additionally, since $\mathbf{Q}_d(t_i)$ is calculated from $\mathbf{\Phi}(t_i, \tau)$ using

$$\mathbf{Q}_d(t_{i-1}) = \int_{t_{i-1}}^{t_i} \mathbf{\Phi}(t_i, \tau)\, \mathbf{G}(\tau)\, \mathbf{Q}(\tau)\, \mathbf{G}^T(\tau)\, \mathbf{\Phi}^T(t_i, \tau)\, d\tau \tag{3.19}$$

this scenario is related to the unknown $\mathbf{Q}_d(t_i)$ scenario of Section 3.3.2. Table 3.3 lists the relationship of the key elemental filter matrices and vectors for this scenario. Under the conditions in Table 3.3, specifically each filter operating on the same measurement vector and the transformation $\mathbf{T}_j(t_i)$ varying between the elemental filters, the derivation of the mean and covariance for $\mathbf{r}_j^*(t_i)$ is exactly the same as in Section 3.3.2. Thus, Equation 3.15 applies to this scenario.

3.3.4 Residual Properties for Varying $\mathbf{R}(t_i)$ Parameter. A third parameter estimatable using an MMAE is the measurement noise covariance matrix $\mathbf{R}(t_i)$. In this problem, multiple measurement noise covariance matrices are hypothesized in the elemental filters. The MMAE determines which elemental filter contains the correct $\mathbf{R}(t_i)$ matrix. Such an application arises in failure detection if the noise in a sensor changes as the result of a failure (or partial failure). This application also arises when hypothesizing differ-

Table 3.4 Relationship Among Elemental Filter Matrices and Vectors for Unknown $\mathbf{R}(t_i)$ Application

Same in All Elemental Filters	**Different in Each Elemental Filter**
$\mathbf{\Phi}(t_i, t_{i-1})$	$\widehat{\mathbf{x}}(t_i^-)$
$\mathbf{B}_d(t_i)$	$\widehat{\mathbf{x}}(t_i^+)$
$\mathbf{u}(t_i)$	$\mathbf{P}(t_i^-)$
$\mathbf{H}(t_i)$	$\mathbf{P}(t_i^+)$
$\mathbf{Q}_d(t_i)$	$\mathbf{K}(t_i)$
$\mathbf{z}(t_i)$	$\mathbf{r}^*(t_i)$
	$\mathbf{A}^*(t_i)$
	$\mathbf{T}(t_i)$
	$\mathbf{R}(t_i)$

ent levels of interference or jamming in GPS. Table 3.4 shows the relationship of the key Kalman filter parameters among the elemental filters. Under the conditions in Table 3.4, specifically each filter operating on the same measurement vector and the transformation matrix $\mathbf{T}_j(t_i)$ varying between the elemental filters, the derivation of the mean and covariance for $\mathbf{r}_j^*(t_i)$ is exactly the same as in Section 3.3.2 and Equation (3.15) applies to this scenario as well.

3.4 Effect of $\mathbf{T}(t_i)$ Transformation

The transformation $\mathbf{T}(t_i) = \left[\mathbf{I} - (1-\gamma)\mathbf{H}(t_i)\mathbf{K}(t_i)\right]$ occurs frequently in derivations using $\mathbf{r}^*(t_i)$. To understand the effect of $\mathbf{T}(t_i)$ on the traditional residual $\mathbf{r}(t_i^-)$, define the difference $\mathbf{\Delta r}(t_i)$ as:

$$\mathbf{\Delta r}(t_i) \stackrel{\triangle}{=} \mathbf{r}(t_i^-) - \mathbf{r}^*(t_i) \tag{3.20}$$

Using Equation (3.2), it can be shown that

$$\mathbf{\Delta r}(t_i) = (1-\gamma)\mathbf{H}(t_i)\mathbf{K}(t_i)\mathbf{r}(t_i^-) \tag{3.21}$$

Examination of the Kalman filter state update equation from Table 2.1, repeated here for convenience,

$$\widehat{\mathbf{x}}\left(t_i^+\right) \;=\; \widehat{\mathbf{x}}\left(t_i^-\right) \;+\; \mathbf{K}\left(t_i\right)\left[\mathbf{z}_i \;-\; \mathbf{H}\left(t_i\right)\widehat{\mathbf{x}}\left(t_i^-\right)\right]$$

reveals that the Kalman gain matrix $\mathbf{K}\left(t_i\right)$ transforms measurement domain vectors into the state space. It also shows that the measurement matrix $\mathbf{H}\left(t_i\right)$ transforms state vectors into the measurement domain. Thus, multiplication by $\mathbf{H}\left(t_i\right)\mathbf{K}\left(t_i\right)$ transforms the measurement domain residual vector into the state space, then back into the measurement space. A two dimensional representation is shown in Figure 3.1. The scalar $(1 \;-\; \gamma)$ is thought of as a gain which determines how much the traditional residual vector is scaled. After scaling by $(1 \;-\; \gamma)$, the residual vector is rotated (and possibly scaled further) by the $\mathbf{H}\left(t_i\right)\mathbf{K}\left(t_i\right)$ multiplication. In the extreme case of $\gamma \;=\; 1$ (i.e., $\mathbf{r}^*\left(t_i\right) = \mathbf{r}\left(t_i^-\right)$), the scalar gain is 0, and $\Delta\mathbf{r}\left(t_i\right)$ is $\mathbf{0}$. In general, $\mathbf{T}\left(t_i\right)$ is a transformation, not a projection, since $\mathbf{T}\left(t_i\right)$ is not idempotent. This fact was confirmed using a simulation to calculate $\mathbf{T}\left(t_i\right)$ in a measurement bias application. Although a formal proof that $\mathbf{T}\left(t_i\right)$ is never a projection was not pursued, it is expected that $\mathbf{T}\left(t_i\right)$ will not be a projection except possibly in very unusual cases.

3.5 Hypothesis Conditional Probability Derivation Using Generalized Residuals

This section is similar to Section 2.3, deriving the probability that each GRMMAE elemental filter is correct, based on a linear combination of the traditional Kalman filter residual and the post-fit residual. To calculate the probability that a given elemental filter contains the correct parameter $\mathbf{a}$, consider the conditional probability density function

$$f_{\mathbf{a}|\mathbf{r}^*(t_i),\,\mathbf{Z}(t_{i-1})}\left(\boldsymbol{\alpha}|\mathbf{r}^*\left(t_i\right) = \boldsymbol{\rho}_i^*, \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\right) \tag{3.22}$$

This is equivalent in information content to $f_{\mathbf{z}(t_i)|\mathbf{a},\mathbf{Z}(t_{i-1})}\left(\boldsymbol{\zeta}_i|\mathbf{a} = \boldsymbol{\alpha}, \mathbf{Z}\left(t_{i-1}\right) = \mathbf{Z}_{i-1}\right)$ from Equation (2.37). Equation (3.22) is used to derive a hypothesis conditional probability formula similar to Equation (2.41), but acting upon the generalized residual, rather than the

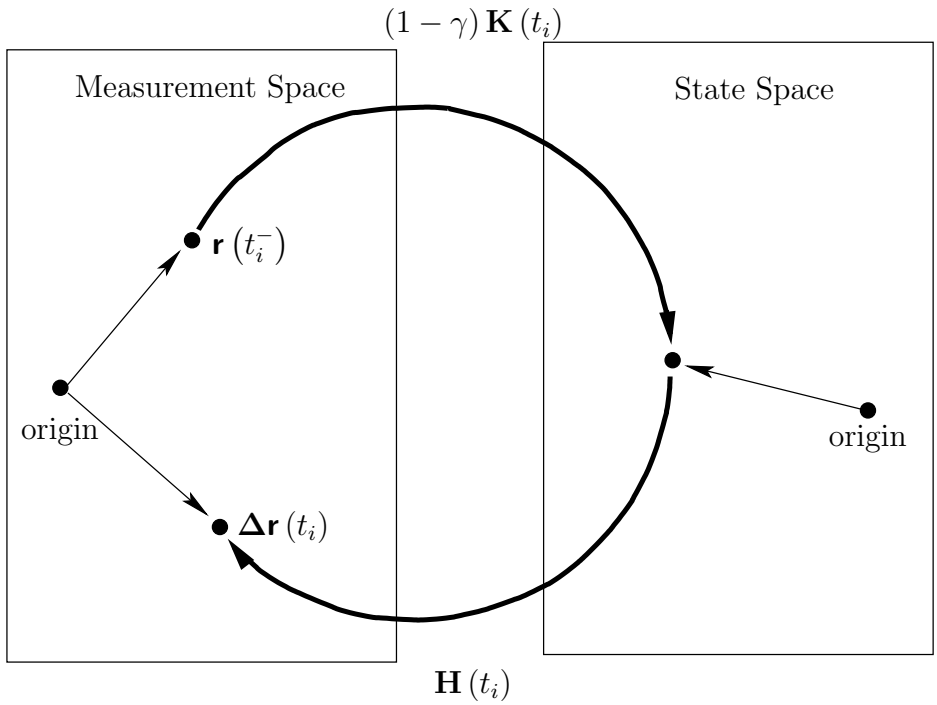


Figure 3.1 2-D Representation of Difference Between Traditional and Generalized Residual Vectors

traditional residual. Similar to Section 2.3, assume $\mathbf{a}$ can take any value in the continuous range $\mathbf{A} \subset \mathbb{R}^p$. Applying Baye's rule to Equation (3.22) yields

$$f_{\mathbf{a}|\mathbf{r}^*(t_i),\,\mathbf{Z}(t_{i-1})}\left(\boldsymbol{\alpha}|\mathbf{r}^*\left(t_i\right)=\boldsymbol{\rho}_i^*,\mathbf{Z}\left(t_{i-1}\right)=\boldsymbol{\mathcal{Z}}_{i-1}\right) \;=\; \frac{f_{\mathbf{r}^*(t_i)|\mathbf{a},\,\mathbf{Z}(t_{i-1})}\left(\boldsymbol{\rho}^*|\mathbf{a}=\boldsymbol{\alpha},\mathbf{Z}\left(t_{i-1}\right)=\mathbf{Z}_{i-1}\right) f_{\mathbf{a}|\mathbf{Z}(t_{i-1})}\left(\boldsymbol{\alpha}|\mathbf{Z}\left(t_{i-1}\right)=\mathbf{Z}_{i-1}\right)}{\int\limits_{\mathbf{A}} f_{\mathbf{r}^*(t_i)|\mathbf{a},\,\mathbf{Z}(t_{i-1})}\left(\boldsymbol{\rho}^*|\mathbf{a}=\boldsymbol{\alpha},\mathbf{Z}\left(t_{i-1}\right)=\mathbf{Z}_{i-1}\right) f_{\mathbf{a}|\mathbf{Z}(t_{i-1})}\left(\boldsymbol{\alpha}|\mathbf{Z}\left(t_{i-1}\right)=\mathbf{Z}_{i-1}\right)\mathbf{d}\boldsymbol{\alpha}} \tag{3.23}$$

Next, for feasible implementation, assume the parameter vector is restricted to be a member of the discrete set $\{\mathbf{a}_1,\ \mathbf{a}_2,\ \ldots\ ,\mathbf{a}_K\}$. Additionally, assume that the correct parameter value is an element of that discrete set. Then, define the hypothesis conditional probability as

$$p_k^*\left(t_i\right) \;\triangleq\; \text{prob}\left\{\mathbf{a} \;=\; \mathbf{a}_k|\mathbf{r}^*\left(t_i\right)=\mathbf{r}_k^*\left(t_i\right),\ \mathbf{Z}\left(t_{i-1}\right)=\mathbf{Z}_{i-1}\right\} \tag{3.24}$$

Under these conditions, the conditional hypothesis probability can be calculated by evaluating Equation (3.23) using the realization of the measurement history $\mathbf{Z}_{i-1}$.

$$\begin{aligned} p_k^*(t_i) &= \frac{f_{\mathbf{r}^*(t_i)|\mathbf{a},\ \mathbf{Z}(t_{i-1})}\left(\mathbf{r}_k^*(t_i)\,|\mathbf{a}=\mathbf{a}_k, \mathbf{Z}(t_{i-1})=\mathbf{Z}_{i-1}\right) p_{\mathbf{a}_k|\mathbf{Z_{i-1}}}}{\int\limits_{\mathbf{A}} f_{\mathbf{r}^*(t_i)|\mathbf{a},\ \mathbf{Z}(t_{i-1})}\left(\mathbf{r}^*(t_i)\,|\mathbf{a}=\boldsymbol{\alpha}, \mathbf{Z}(t_{i-1})=\mathbf{Z}_{i-1}\right) \sum\limits_{j=1}^{K} p_{\mathbf{a}|\mathbf{Z_{i-1}}}\delta(\boldsymbol{\alpha} - \mathbf{a}_j)\,\mathbf{d}\boldsymbol{\alpha}} \\ &= \frac{f_{\mathbf{r}^*(t_i)|\mathbf{a},\ \mathbf{Z}(t_{i-1})}\left(\mathbf{r}_k^*(t_i)\,|\mathbf{a}=\mathbf{a}_k, \mathbf{Z}(t_{i-1})=\mathbf{Z}_{i-1}\right) p_{\mathbf{a}_k|\mathbf{Z_{i-1}}}}{\sum\limits_{j=1}^{K}\int\limits_{\mathbf{A}} f_{\mathbf{r}^*(t_i)|\mathbf{a},\ \mathbf{Z}(t_{i-1})}\left(\mathbf{r}^*(t_i)\,|\mathbf{a}=\boldsymbol{\alpha}, \mathbf{Z}(t_{i-1})=\mathbf{Z}_{i-1}\right) p_{\mathbf{a}|\mathbf{Z_{i-1}}}\delta(\boldsymbol{\alpha} - \mathbf{a}_j)\,\mathbf{d}\boldsymbol{\alpha}} \\ p_k^*(t_i) &= \frac{f_{\mathbf{r}^*(t_i)|\mathbf{a},\ \mathbf{Z}(t_{i-1})}\left(\mathbf{r}_k^*(t_i)\,|\mathbf{a}=\mathbf{a}_k, \mathbf{Z}(t_{i-1})=\mathbf{Z}_{i-1}\right) p_{\mathbf{a}_k|\mathbf{Z_{i-1}}}}{\sum\limits_{j=1}^{K} f_{\mathbf{r}^*(t_i)|\mathbf{a},\ \mathbf{Z}(t_{i-1})}\left(\mathbf{r}_j^*(t_i)\,|\mathbf{a}=\mathbf{a}_j, \mathbf{Z}(t_{i-1})=\mathbf{Z}_{i-1}\right) p_{\mathbf{a}_j|\mathbf{Z_{i-1}}}} \end{aligned} \tag{3.25}$$

where $p_{\mathbf{a}_k|\mathbf{Z}_{i-1}}$ is the discrete probability mass associated with the parameter realization $\mathbf{a}_k$ given the measurement history realization $\mathbf{Z}_{i-1}$. To see that $p_{\mathbf{a}_k|\mathbf{Z}_{i-1}} = p_k(t_{i-1})$, consider

$$\begin{aligned} p_{\mathbf{a}_k|\mathbf{Z}_{i-1}} &= p_{\mathbf{a}_k|\mathbf{z}_{i-1},\ \mathbf{Z}_{i-2}} \\ &= p_{\mathbf{a}_k|\mathbf{z}_{i-1}-[\mathbf{I}\ -\ (1\ -\ \gamma)\mathbf{H}_k(t_i)\mathbf{K}_k(t_i)]\mathbf{r}_k\left(t_{i-1}^-\right),\ \mathbf{Z}_{i-2}} \\ &= p_{\mathbf{a}_k|\mathbf{r}_k^*(t_{i-1}),\ \mathbf{Z}_{i-2}} \\ p_{\mathbf{a}_k|\mathbf{Z}_{i-1}} &= p_k^*(t_{i-1}) \end{aligned} \tag{3.26}$$

Substituting Equation (3.26) into Equation (3.25) yields the final recursive relationship

$$p_k^*(t_i) = \frac{f_{\mathbf{r}^*(t_i)|\mathbf{a},\ \mathbf{Z}(t_{i-1})}\left(\mathbf{r}_k^*(t_i)\,|\mathbf{a}=\mathbf{a}_k, \mathbf{Z}(t_{i-1})=\mathbf{Z}_{i-1}\right) p_k^*(t_{i-1})}{\sum\limits_{j=1}^{K} f_{\mathbf{r}^*(t_i)|\mathbf{a},\ \mathbf{Z}(t_{i-1})}\left(\mathbf{r}_j^*(t_i)\,|\mathbf{a}=\mathbf{a}_j, \mathbf{Z}(t_{i-1})=\mathbf{Z}_{i-1}\right) p_j^*(t_{i-1})} \tag{3.27}$$

In Equation (3.27), the density $f_{\mathbf{r}^*(t_i)|\mathbf{a},\ \mathbf{Z}(t_{i-1})}\left(\boldsymbol{\rho}^*(t_i)\,|\mathbf{a}=\boldsymbol{\alpha},\ \mathbf{Z}(t_{i-1})=\boldsymbol{\mathcal{Z}}_{i-1}\right)$ is Gaussian. To demonstrate this, consider $f_{\mathbf{z}(t_i)|\mathbf{a},\ \mathbf{Z}(t_{i-1})}\left(\boldsymbol{\zeta}_i|\mathbf{a}=\boldsymbol{\alpha},\ \mathbf{Z}(t_{i-1})=\boldsymbol{\mathcal{Z}}_{i-1}\right)$. From Section 2.3 and [30], this second density is Gaussian. As shown in Equation (3.2), the gen-

eralized residual $\mathbf{r}^*(t_i)$ is a linear function of $\mathbf{z}(t_i)$ and $\widehat{\mathbf{x}}(t_i^-)$. When Equation (3.1) is conditioned on $\mathbf{Z}(t_{i-1}) = \mathbf{Z}_{i-1}$, as in Equation (3.27), $\widehat{\mathbf{x}}(t_i^-)$ becomes the realization $\widehat{\mathbf{x}}(t_i^-)$. Under this condition, $\mathbf{r}^*(t_i)$ is a linear function of the random variable $\mathbf{z}(t_i)$, and the realization $\widehat{\mathbf{x}}(t_i^-)$. Therefore, since $f_{\mathbf{z}(t_i)|\mathbf{a},\ \mathbf{Z}(t_{i-1})}(\boldsymbol{\zeta}_i|\mathbf{a}=\boldsymbol{\alpha},\ \mathbf{Z}(t_{i-1})=\boldsymbol{\mathcal{Z}}_{i-1})$ is Gaussian [30], $f_{\mathbf{r}^*(t_i)|\mathbf{a},\ \mathbf{Z}(t_{i-1})}(\boldsymbol{\rho}_k^*|\mathbf{a}=\boldsymbol{\alpha},\ \mathbf{Z}(t_{i-1})=\boldsymbol{\mathcal{Z}}_{i-1})$ is also Gaussian. The mean and covariance of $\mathbf{r}^*(t_i)$, derived in Section 3.1, completely specify this conditional density.

3.6 Relationship Between $p_k(t_i)$ and $p_k^*(t_i)$

The traditional MMAE ($\gamma = 1$) calculates elemental filter conditional probabilities from Equation (2.41), repeated here for convenience:

$$p_k(t_i) = \frac{f_{\mathbf{z}(t_i)|\mathbf{a},\ \mathbf{Z}(t_{i-1})}(\mathbf{z}_i|\mathbf{a}=\mathbf{a}_k, \mathbf{Z}(t_{i-1})=\mathbf{Z}_{i-1})\, p_k(t_{i-1})}{\sum_{j=1}^{K} f_{\mathbf{z}(t_i)|\mathbf{a},\ \mathbf{Z}(t_{i-1})}(\mathbf{z}_i|\mathbf{a}=\mathbf{a}_j, \mathbf{Z}(t_{i-1})=\mathbf{Z}_{i-1})\, p_j(t_{i-1})} \tag{2.41}$$

Equation (3.27) and Equation (2.41) must evaluate to the same value when $\gamma = 1$ (i.e., the generalized residual becomes the traditional residual). Therefore, there is a relationship between these two equations.

Section 3.2 shows that $\mathbf{r}_k^*(t_i)$ is Gaussian and zero-mean, with covariance given by Equation (3.4) (assuming the k^{th} elemental filter contains the correct parameter). Thus, the conditional density is

$$\begin{aligned} &f_{\mathbf{r}_k^*(t_i)|\mathbf{a},\ \mathbf{Z}(t_{i-1})}(\mathbf{r}_k^*(t_i)|\mathbf{a}=\mathbf{a}_k, \mathbf{Z}(t_{i-1})=\mathbf{Z}_{i-1}) \\ &\quad = \frac{1}{(2\pi)^{\frac{m}{2}}|\mathbf{A}_k^*(t_i)|^{1/2}} \exp\Big\{-\frac{1}{2}\mathbf{r}_k^{*T}(t_i)\,\mathbf{A}_k^{*-1}(t_i)\,\mathbf{r}_k^*(t_i)\Big\} \\ &\quad = \beta_k^*(t_i) \exp\Big\{-\frac{1}{2}\mathbf{r}_k^{*T}(t_i)\Big[\mathbf{T}_k(t_i)\,\mathbf{A}_k^{-1}(t_i^-)\,\mathbf{T}_k^T(t_i)\Big]^{-1}\mathbf{r}_k^*(t_i)\Big\} \end{aligned} \tag{3.28}$$

In Equation (3.28), $\beta_k^*(t_i)$, $\mathbf{A}_k(t_i^-)$ and $\mathbf{T}_k(t_i)$ are given by

$$
\begin{aligned}
\beta_k^*(t_i) &= \frac{1}{(2\pi)^{\frac{m}{2}} \left|\mathbf{T}_k(t_i)\left[\mathbf{H}_k(t_i)\mathbf{P}_k\left(t_i^-\right)\mathbf{H}_k^T(t_i) + \mathbf{R}_k(t_i)\right]\mathbf{T}_k^T(t_i)\right|^{\frac{1}{2}}} \\
\mathbf{A}_k\left(t_i^-\right) &= \left[\mathbf{H}_k(t_i)\mathbf{P}_k\left(t_i^-\right)\mathbf{H}_k^T(t_i) + \mathbf{R}_k(t_i)\right] \\
\mathbf{T}_k(t_i) &= \left[\mathbf{I} - (1-\gamma)\mathbf{H}_k(t_i)\mathbf{K}_k(t_i)\right]
\end{aligned} \tag{3.29}
$$

Consider the term inside the exponential first

$$
\begin{aligned}
&-\frac{1}{2}\mathbf{r}_k^{*^T}(t_i)\,\mathbf{A}_k^{*^{-1}}(t_i)\,\mathbf{r}_k^*(t_i) \\
&= -\frac{1}{2}\mathbf{r}_k^{*^T}(t_i)\left[\mathbf{T}_k(t_i)\,\mathbf{A}_k^{-1}\left(t_i^-\right)\mathbf{T}_k^T(t_i)\right]^{-1}\mathbf{r}_k^*(t_i) \\
&= -\frac{1}{2}\mathbf{r}_k^T\left(t_i^-\right)\mathbf{T}_k^T(t_i)\left[\mathbf{T}_k(t_i)\left[\mathbf{H}_k(t_i)\mathbf{P}_k\left(t_i^-\right)\mathbf{H}_k^T(t_i) + \mathbf{R}_k(t_i)\right]\mathbf{T}_k^T(t_i)\right]^{-1}\mathbf{T}_k(t_i)\,\mathbf{r}_k\left(t_i^-\right) \\
&= -\frac{1}{2}\mathbf{r}_k^T\left(t_i^-\right)\mathbf{T}_k^T(t_i)\left[\mathbf{T}_k^T(t_i)\right]^{-1}\left[\mathbf{H}_k(t_i)\mathbf{P}_k\left(t_i^-\right)\mathbf{H}_k^T(t_i) + \mathbf{R}_k(t_i)\right]^{-1}\mathbf{T}_k^{-1}(t_i)\,\mathbf{T}_k(t_i)\,\mathbf{r}_k\left(t_i^-\right) \\
&= -\frac{1}{2}\mathbf{r}_k^T\left(t_i^-\right)\left[\mathbf{H}_k(t_i)\mathbf{P}_k\left(t_i^-\right)\mathbf{H}_k^T(t_i) + \mathbf{R}_k(t_i)\right]^{-1}\mathbf{r}_k\left(t_i^-\right) \\
&= -\frac{1}{2}\mathbf{r}_k^T\left(t_i^-\right)\mathbf{A}_k^{-1}\left(t_i^-\right)\mathbf{r}_k\left(t_i^-\right)
\end{aligned} \tag{3.30}
$$

Comparing Equation (3.30) to Equation (2.47) shows that the term inside the exponential for $f_{\mathbf{r}_k^*(t_i)|\mathbf{a},\,\mathbf{Z}(t_{i-1})}\left(\mathbf{r}_k^*(t_i)\,|\mathbf{a}=\mathbf{a}_k,\mathbf{Z}(t_{i-1})=\mathbf{Z}_{i-1}\right)$ is exactly the same as the term inside the exponential for $f_{\mathbf{z}(t_i)|\mathbf{a},\,\mathbf{Z}(t_{i-1})}\left(\mathbf{z}_i|\mathbf{a}=\mathbf{a}_k,\mathbf{Z}(t_{i-1})=\mathbf{Z}_{i-1}\right)$. This implies that the GRMMAE design parameter γ has no effect on the likelihood quotients. Thus, regardless of the value of γ, the likelihood quotient inside the Gaussian densities of Equation (3.27) is $-1/2\mathbf{r}_k^T(t_i)\,\mathbf{A}_k^{-1}(t_i)\,\mathbf{r}_k(t_i)$, the same as it is in the traditional MMAE.

Now, consider the $\beta_k^*(t_i)$ term in front of the exponential in Equation (3.28). The determinant $\left|\mathbf{T}_k(t_i)\left[\mathbf{H}_k(t_i)\mathbf{P}_k\left(t_i^-\right)\mathbf{H}_k^T(t_i) + \mathbf{R}(t_i)\right]\mathbf{T}_k^T(t_i)\right|$ can be broken into the 3 separate determinants in the first line of Equation (3.31).

$$
\beta_k^*(t_i) = \frac{1}{(2\pi)^{\frac{m}{2}}\left\{\left|\mathbf{T}_k(t_i)\right|\left|\left[\mathbf{H}_k(t_i)\mathbf{P}_k\left(t_i^-\right)\mathbf{H}_k^T(t_i) + \mathbf{R}_k(t_i)\right]\right|\left|\mathbf{T}_k^T(t_i)\right|\right\}^{\frac{1}{2}}}
$$

$$
\begin{aligned}
\beta_k^*(t_i) &= \frac{1}{(2\pi)^{\frac{m}{2}} \left\{ |\mathbf{T}_k(t_i)| \, |\left[\mathbf{H}_k(t_i)\mathbf{P}_k\left(t_i^-\right)\mathbf{H}_k^T(t_i) \; + \; \mathbf{R}_k(t_i)\right]| \, |\mathbf{T}_k(t_i)| \right\}^{\frac{1}{2}}} \\
&= \frac{1}{(2\pi)^{\frac{m}{2}} \operatorname{abs}\left(|\mathbf{T}_k(t_i)|\right) \left\{ \, |\left[\mathbf{H}_k(t_i)\mathbf{P}_k\left(t_i^-\right)\mathbf{H}_k^T(t_i) \; + \; \mathbf{R}_k(t_i)\right]| \, \right\}^{\frac{1}{2}}} \\
\beta_k^*(t_i) &= \operatorname{abs}\left(|\mathbf{T}_k(t_i)|\right)^{-1} \frac{1}{(2\pi)^{\frac{m}{2}} \left\{ \, |\left[\mathbf{H}_k(t_i)\mathbf{P}_k\left(t_i^-\right)\mathbf{H}_k^T(t_i) \; + \; \mathbf{R}_k(t_i)\right]| \, \right\}^{\frac{1}{2}}}
\end{aligned} \tag{3.31}
$$

In Equation (3.31) the determinant $|\mathbf{T}_k(t_i)|$ is squared inside the square root, then brought outside the square root. When it is brought outside, the positive square root is used. Therefore, the notation $\operatorname{abs}\left(|\mathbf{T}_k(t_i)|\right)$ is used to denote a positive square root.

Equation (3.31) equals the beta term from $f_{\mathbf{z}(t_i)|\mathbf{a},\ \mathbf{Z_{i-1}}}(\mathbf{z}_i|\mathbf{a}=\mathbf{a}_k, \mathbf{Z}(t_{i-1})=\mathbf{Z}_{i-1})$, in Equation (2.47), multiplied by the scalar term $\operatorname{abs}\left(|\mathbf{T}_k(t_i)|^{-1}\right)$. Thus,

$$
\begin{aligned}
&f_{\mathbf{r}_k^*(t_i)|\mathbf{a},\ \mathbf{Z}(t_{i-1})}(\mathbf{r}_k^*(t_i)|\mathbf{a}=\mathbf{a}_k,\ \mathbf{Z}(t_{i-1})=\mathbf{Z}_{i-1}) \\
&\qquad = \operatorname{abs}\left(|\mathbf{T}_k(t_i)|\right)^{-1} f_{\mathbf{z}(t_i)|\mathbf{a},\ \mathbf{Z}(t_{i-1})}(\mathbf{z}(t_i)|\mathbf{a}=\mathbf{a}_k,\ \mathbf{Z}(t_{i-1})=\mathbf{Z}_{i-1})
\end{aligned} \tag{3.32}
$$

Under the condition $\gamma = 1$ (i.e., $\mathbf{r}_k^*(t_i) = \mathbf{r}_k\left(t_i^-\right)$), $\mathbf{T}_k(t_i) = \mathbf{I}$ and

$$
\begin{aligned}
&f_{\mathbf{r}_k^*(t_i)|\mathbf{a},\mathbf{Z}(t_{i-1})}(\mathbf{r}_k^*(t_i)|\mathbf{a}=\mathbf{a}_k, \mathbf{Z}(t_{i-1})=\mathbf{Z}_{i-1}) \\
&\qquad = f_{\mathbf{z}(t_i)|\mathbf{a},\mathbf{Z}(t_{i-1})}(\mathbf{z}(t_i)|\mathbf{a}=\mathbf{a}_k, \mathbf{Z}(t_{i-1})=\mathbf{Z}_{i-1})
\end{aligned} \tag{3.33}
$$

Thus, when $\gamma = 1$, the GRMMAE is equivalent to a traditional MMAE.

Equations (3.30) through (3.32) demonstrate that the GRMMAE design parameter γ affects only the $\beta^*(t_i)$ normalization term in front of the exponential in the Gaussian densities of Equation (3.27). Thus, γ in the generalized residual has the effect of artificially increasing or decreasing the $\beta^*(t_i)$ term magnitude in these Gaussian densities. Chapter IV demonstrates that this results in a susceptibility to the β-dominance effect.

3.7 Optimization of $p_k^*(t_i)$

The objective of the γ parameter in the generalized residual is to improve the rate of convergence to the correct elemental filter of the GRMMAE compared to the traditional MMAE. Thus, under the assumption that the k^{th} elemental filter is correct, the optimal value of γ should maximize $p_k^*(t_i)$ given in Equation (3.27). All GRMMAE applications can be divided into 2 groups to accomplish this objective. The first group consists of applications in which the transformation matrix $\mathbf{T}(t_i)$ is identical in all elemental filters. For this scenario, any γ may be used with identical results achieved. This will be demonstrated in Section 3.7.1. The second group consists of applications in which the transformation matrix $\mathbf{T}(t_i)$ varies among the elemental filters. In these applications, a specific γ value will maximize $p_k^*(t_i)$. The specific optimal value of γ will vary dependent upon the application. However, these applications will be susceptible to β-dominance if $\gamma \neq 1$, as demonstrated in Sections 3.7.2 and 4.3.

In both cases, the same general procedure is used to determine the optimal γ value:

1. Determine the generalized residual $\mathbf{r}_j^*(t_i)$ in terms of the traditional residual for the correct elemental filter $\mathbf{r}_k\left(t_i^-\right)$.
2. Determine the covariance of the generalized residual $\mathbf{A}_j^*(t_i)$ in terms of the covariance of the traditional residual for the correct elemental filter $\mathbf{A}_k\left(t_i^-\right)$.
3. Perform algebraic substitution and simplification to express the generalized hypothesis conditional probability of Equation (3.27) in terms of $\mathbf{r}_k\left(t_i^-\right)$ and $\mathbf{A}_k\left(t_i^-\right)$.
4. Determine the γ value which maximizes $p_k^*(t_i)$ (either in steady state or at each measurement update).

Steps 1 and 2 have already been accomplished in Sections 3.2 and 3.3. Because of the β-dominance effect, Step 4 will not be pursued. Rather, the designer is strongly cautioned against using $\gamma \neq 1$ unless $\mathbf{T}(t_i)$ is identical in all elemental filters. The reason for this caution will be demonstrated in Sections 3.7.2 and 4.3.

3.7.1 Same $\mathbf{T}(t_i)$ In All Elemental Filters. In this section, Equation (3.27) is applied to applications with $\mathbf{T}(t_i)$ identical in all elemental filters. Significantly general

notation and equations applicable to all cases with identical $\mathbf{T}(t_i)$ in all elemental filters cannot be developed. Therefore, the specific example of estimating a constant measurement bias, as detailed in Section 3.3.1, is shown. This will demonstrate the 4-step procedure outlined in the previous section. Other applications with the same $\mathbf{T}(t_i)$ in all elemental filters can be similarly derived. Begin with Equation (3.27), substituting the residuals from the elemental filters into the densities

$$p_k^*(t_i) = \frac{\frac{1}{(2\pi)^{m/2}\text{abs}\left(|\mathbf{T}(t_i)|\right)|\mathbf{A}_k\left(t_i^-\right)|^{1/2}} \exp\left\{-\frac{1}{2}\mathbf{r}_k^T\left(t_i^-\right)\mathbf{A}_k^{-1}\left(t_i^-\right)\mathbf{r}_k\left(t_i^-\right)\right\} p_k^*(t_{i-1})}{\sum_{j=1}^{K}\frac{1}{(2\pi)^{m/2}|\mathbf{A}_j^*(t_i)|^{1/2}} \exp\left\{-\frac{1}{2}\mathbf{r}_j^{*T}(t_i)\mathbf{A}_j^{*-1}(t_i)\mathbf{r}_j^*(t_i)\right\} p_j^*(t_{i-1})} \tag{3.34}$$

where the numerator comes from Equation (3.32)

Next, use Equation (3.13) to substitute for $\mathbf{A}_j^*(t_i)$

$$p_k^*(t_i) = \frac{\frac{1}{(2\pi)^{m/2}\text{abs}\left(|\mathbf{T}(t_i)|\right)|\mathbf{A}_k\left(t_i^-\right)|^{1/2}} \exp\left\{-\frac{1}{2}\mathbf{r}_k^T\left(t_i^-\right)\mathbf{A}_k^{-1}\left(t_i^-\right)\mathbf{r}_k\left(t_i^-\right)\right\} p_k^*(t_{i-1})}{\sum_{j=1}^{K}\frac{1}{(2\pi)^{m/2}\text{abs}\left(|\mathbf{T}(t_i)|\right)|\mathbf{A}_k\left(t_i^-\right)|^{1/2}} \exp\left\{-\frac{1}{2}\mathbf{r}_j^{*T}(t_i)\left[\mathbf{T}(t_i)\mathbf{A}_k\left(t_i^-\right)\mathbf{T}^T(t_i)\right]^{-1}\mathbf{r}_j^*(t_i)\right\} p_j^*(t_{i-1})} \tag{3.35}$$

Notice in Equation (3.35) that all densities in the both the numerator and denominator are multiplied by $1/\left[(2\pi)^{m/2}\,\text{abs}\left(|\mathbf{T}(t_i)|\right)|\mathbf{A}_k\left(t_i^-\right)|^{1/2}\right]$. Therefore, this term is eliminated

$$p_k^*(t_i) = \frac{\exp\left\{-\frac{1}{2}\mathbf{r}_k^T\left(t_i^-\right)\mathbf{A}_k^{-1}\left(t_i^-\right)\mathbf{r}_k\left(t_i^-\right)\right\} p_k^*(t_{i-1})}{\sum_{j=1}^{K}\exp\left\{q_j^*\right\} p_j^*(t_{i-1})} \tag{3.36}$$

where

$$
\begin{aligned}
q_j^* &= -\frac{1}{2}\left[\mathbf{r}_k\left(t_i^-\right)+\mathbf{H}\left(t_i\right)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right)-\widehat{\mathbf{x}}_j\left(t_i^-\right)\right)+\boldsymbol{\Delta}\mathbf{n}_{kj}\right]^T\mathbf{T}^T\left(t_i\right)\mathbf{A}_j^{*^{-1}}\left(t_i\right)\mathbf{T}\left(t_i\right) \\
&\quad \left[\mathbf{r}_k\left(t_i^-\right)+\mathbf{H}\left(t_i\right)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right)-\widehat{\mathbf{x}}_j\left(t_i^-\right)\right)+\boldsymbol{\Delta}\mathbf{n}_{kj}\right] \\
&= -\frac{1}{2}\left[\mathbf{r}_k\left(t_i^-\right)+\mathbf{H}\left(t_i\right)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right)-\widehat{\mathbf{x}}_j\left(t_i^-\right)\right)+\boldsymbol{\Delta}\mathbf{n}_{kj}\right]^T \\
&\quad \mathbf{T}^T\left(t_i\right)\left[\mathbf{T}\left(t_i\right)\mathbf{A}_k\left(t_i^-\right)\mathbf{T}^T\left(t_i\right)\right]^{-1}\mathbf{T}\left(t_i\right) \\
&\quad \left[\mathbf{r}_k\left(t_i^-\right)+\mathbf{H}\left(t_i\right)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right)-\widehat{\mathbf{x}}_j\left(t_i^-\right)\right)+\boldsymbol{\Delta}\mathbf{n}_{kj}\right] \\
q_j^* &= -\frac{1}{2}\left[\mathbf{r}_k\left(t_i^-\right)+\mathbf{H}\left(t_i\right)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right)-\widehat{\mathbf{x}}_j\left(t_i^-\right)\right)+\boldsymbol{\Delta}\mathbf{n}_{kj}\right]^T\mathbf{A}_k^{-1}\left(t_i^-\right) \\
&\quad \left[\mathbf{r}_k\left(t_i^-\right)+\mathbf{H}\left(t_i\right)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right)-\widehat{\mathbf{x}}_j\left(t_i^-\right)\right)+\boldsymbol{\Delta}\mathbf{n}_{kj}\right] \qquad \forall\ j
\end{aligned}
\tag{3.37}
$$

Note that $\mathbf{T}\left(t_i\right)$ has been eliminated from Equation (3.37). Since $\mathbf{T}\left(t_i\right)$ is the only variable that depends upon the design parameter γ, $p_k^*\left(t_i\right)$ is independent of γ. This means that Equation (3.35) (or, equivalently, Equation (3.27)) will yield exactly the same hypothesis conditional probability regardless of what generalized residual vector is used in the GRMMAE for the measurement bias determination problem. The γ independence can also be interpreted as meaning that β-stripping, described in Section 2.4.1, is always performed for applications in which $\mathbf{T}\left(t_i\right)$ does not vary among the elemental filters.

3.7.2 Varying $\mathbf{T}\left(t_i\right)$ In Elemental Filters. As in the previous subsection, begin by substituting the residuals from the elemental filters into the densities in Equation (3.27). Assuming that each elemental filter operates on the same measurement, the residual characteristics given in Equation (3.15) hold. Substituting these characteristics in Equation (3.27) yields

$$
p_k^*\left(t_i\right) = \frac{\frac{1}{(2\pi)^{m/2}\mathrm{abs}\left(|\mathbf{T}_k(t_i)|\right)|\mathbf{A}_k\left(t_i^-\right)|^{1/2}}\exp\left\{-\frac{1}{2}\mathbf{r}_k^T\left(t_i^-\right)\mathbf{A}_k^{-1}\left(t_i^-\right)\mathbf{r}_k\left(t_i^-\right)\right\}p_k^*\left(t_{i-1}\right)}{\sum_{j=1}^{K}\frac{1}{(2\pi)^{m/2}\mathrm{abs}\left(|\mathbf{T}_j(t_i)|\right)|\mathbf{A}_k\left(t_i^-\right)|^{1/2}}\exp\left\{q_j^*\left(t_i\right)\right\}p_j^*\left(t_{i-1}\right)} \tag{3.38}
$$

where

$$
\begin{aligned}
q_j^*(t_i) &= -\frac{1}{2}\mathbf{r}_j^{*T}(t_i)\,\mathbf{A}_j^{*^{-1}}(t_i)\,\mathbf{r}_j^*(t_i) \\
&= -\frac{1}{2}\left[\mathbf{r}_k\left(t_i^-\right) + \mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right) - \widehat{\mathbf{x}}_j\left(t_i^-\right)\right)\right]^T \mathbf{T}_j^T(t_i)\,\mathbf{A}_j^{*^{-1}}(t_i)\,\mathbf{T}_j(t_i) \\
&\quad \left[\mathbf{r}_k\left(t_i^-\right) + \mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right) - \widehat{\mathbf{x}}_j\left(t_i^-\right)\right)\right] \\
&= -\frac{1}{2}\left[\mathbf{r}_k\left(t_i^-\right) + \mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right) - \widehat{\mathbf{x}}_j\left(t_i^-\right)\right)\right]^T \mathbf{T}_j^T(t_i)\left[\mathbf{T}_j(t_i)\,\mathbf{A}_k\left(t_i^-\right)\mathbf{T}_j^T(t_i)\right]^{-1} \\
&\quad \mathbf{T}_j(t_i)\left[\mathbf{r}_k\left(t_i^-\right) + \mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right) - \widehat{\mathbf{x}}_j\left(t_i^-\right)\right)\right] \\
q_j^*(t_i) &= -\frac{1}{2}\left[\mathbf{r}_k\left(t_i^-\right) + \mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right) - \widehat{\mathbf{x}}_j\left(t_i^-\right)\right)\right]^T \mathbf{A}_k^{-1}\left(t_i^-\right) \\
&\quad \left[\mathbf{r}_k\left(t_i^-\right) + \mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right) - \widehat{\mathbf{x}}_j\left(t_i^-\right)\right)\right] \qquad \forall\ j
\end{aligned} \tag{3.39}
$$

Next, explicitly remove the component for the k^{th} filter from the summation in the denominator and factor out $1/\left[(2\pi)^{m/2}\,\mathrm{abs}\big(|\mathbf{T}_k(t_i)|\big)|\mathbf{A}_k\left(t_i^-\right)|^{1/2}\right]$ and $p_k(t_{i-1})$ to yield

$$
p_k^*(t_i) = \frac{\exp\{q_k(t_i)\}}{\exp\{q_k(t_i)\} + \sum\limits_{\substack{j=1 \\ j\neq k}}^{K} \mathrm{abs}\left(\frac{|\mathbf{T}_k(t_i)|}{|\mathbf{T}_j(t_i)|}\right)\frac{p_j^*(t_{i-1})}{p_k^*(t_{i-1})}\exp\left\{q_j^*(t_i)\right\}} \tag{3.40}
$$

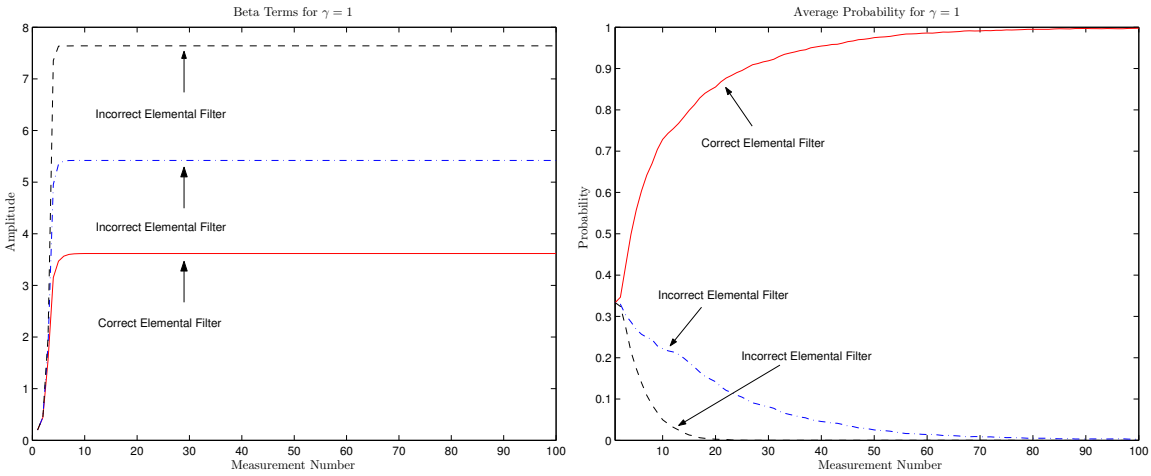
$\mathbf{T}_k(t_i)$, $\mathbf{T}_j(t_i)$, and the prior probabilities are the only terms in Equation (3.40) that depend upon γ. In order to maximize $p_k^*(t_i)$, γ must be chosen to minimize

$$
\mathrm{abs}\left(\frac{|\mathbf{T}_k(t_i)|}{|\mathbf{T}_j(t_i)|}\right)\frac{p_j(t_{i-1})}{p_k(t_{i-1})} \tag{3.41}
$$

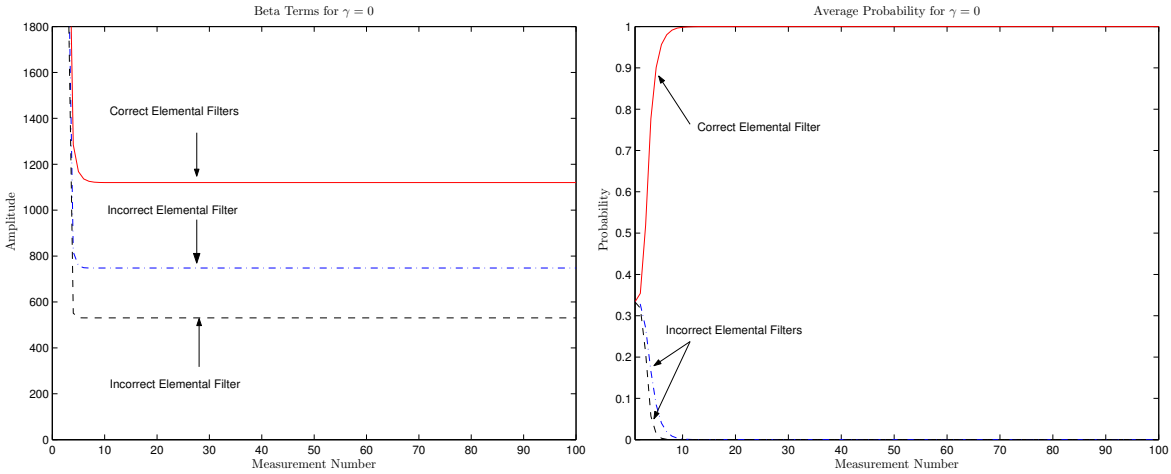
This minimizes the contribution of the incorrect elemental filters, and maximizes $p_k^*(t_i)$. However, setting $\gamma \neq 1$ has potential to cause β-dominance. If the correct filter, k, is known *a priori* (in which case there is no problem to solve), an optimal γ to minimize Equation (3.41) can be found. However, for a given problem, this optimal γ may change if k changes, dependent upon the magnitude of the $\beta_k^*(t_i)$ and $\beta_j^*(t_i)$ terms. Figures 3.2 and 3.3 demonstrate this concept. These figures show the $\beta_j^*(t_i)$ terms (including $j = k$) and mean hypothesis conditional probabilities from 1000 Monte Carlo simulations of an application with varying $\mathbf{T}_j(t_i)$. (The specific models and parameters for the simulations are not presented here because they would distract from the concept being illustrated.) In Figure 3.2, the correct hypothesis is arbitrarily chosen to be the hypothesis with the largest

steady state $\beta_j^*(t_i)$ value for the GRMMAE $\gamma = 0$. For Figure 3.2, a measurement set was generated and used in GRMMAEs with $\gamma = 1$ and $\gamma = 0$ (i.e., equivalent to traditional and post-fit modified MMAEs, respectively). From Figure 3.2, both GRMMAEs converged to the correct hypothesis. However, the GRMMAE with $\gamma = 0$ converged more quickly than the one with $\gamma = 1$. Thus, based on this particular correct hypothesis (i.e., the elemental filter with the largest steady state $\beta_j^*(t_i)$ when $\gamma = 0$), the optimal γ value is 0.

Now, consider Figure 3.3. The same application and system model was used for this figure. However, the true parameter value was changed to the parameter associated with the elemental filter having the smallest steady state $\beta_j^*(t_i)$ values when $\gamma = 0$. Again, a measurement set was generated (associated with this new correct parameter value), and GRMMAEs with $\gamma = 1$ and $\gamma = 0$ were used. From Figure 3.3, the GRMMAE with $\gamma = 1$ (i.e., traditional MMAE equivalent) again converged to the correct hypothesis. However, the GRMMAE with $\gamma = 0$ converged to an incorrect hypothesis. In fact, the GRMMAE with $\gamma = 0$ converged to the same hypothesis as the GRMMAEs in Figure 3.2. This occurred because the pre-computable $\beta_j^*(t_i)$ term (which did not vary between these runs) dominated the hypothesis conditional probability calculation in the GRMMAEs with $\gamma = 0$. For the simulations in Figure 3.3, the optimal γ value is 1. Thus, inconsistent optimal γ values are achieved, dependent upon the relative magnitude of the beta term associated with the correct hypothesis, due to the β-dominance effect. The β-dominance effect is discussed in depth in Section 4.3. However, this illustration demonstrates the β-dominance problem as it relates to the choice of the optimal γ value. Specifically, in some cases, the optimal γ artificially induces β-dominance which causes the GRMMAE to converge to one hypothesis regardless of the realized measurements. Such behavior by the GRMMAE is undesirable because it means that the GRMMAE is ignoring the real world data and is relying entirely on its internal models. Therefore, determining the γ that maximizes $p_k^*(t_i)$ will not be pursued for applications where $\mathbf{T}_j(t_i)$ varies among the elemental filters.

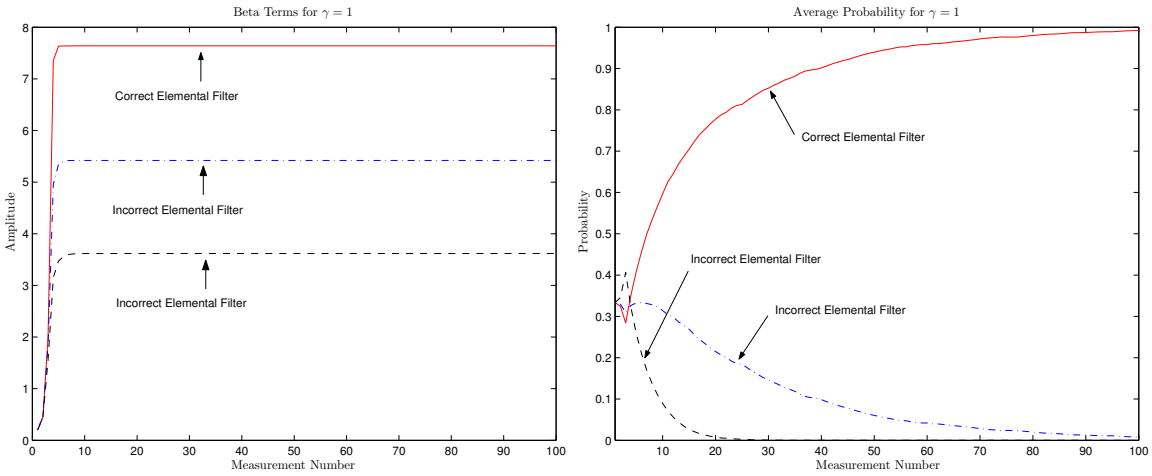


(a) Beta Terms and Average Hypothesis Conditional Probability for GRMMAE with $\gamma = 1$

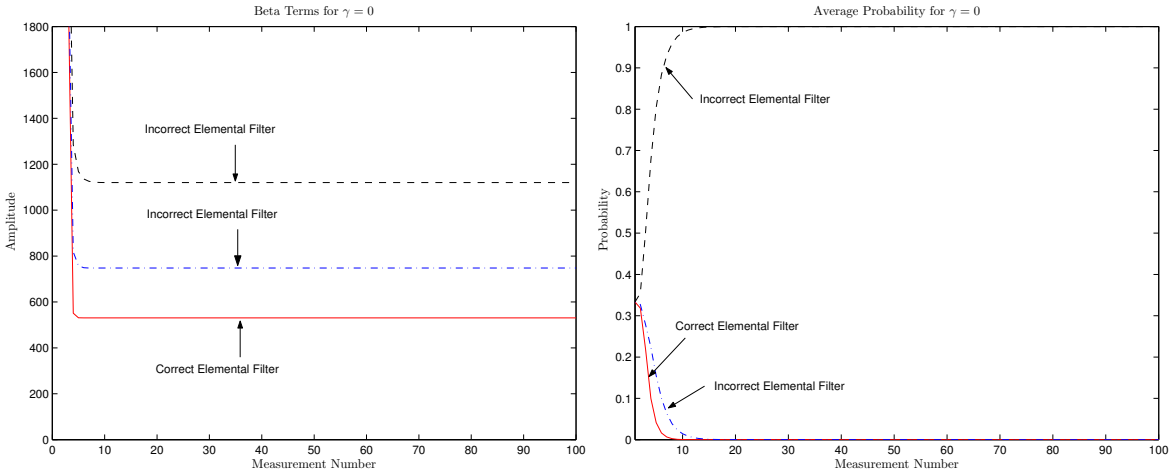


(b) Beta Terms and Average Hypothesis Conditional Probability for GRMMAE with $\gamma = 0$

Figure 3.2 Beta Terms and Average Hypothesis Conditional Probabilities for GRMMAEs with $\gamma = 1$ and $\gamma = 0$ and Correct Hypothesis Having Largest Steady State Beta Term for $\gamma = 0$



(a) Beta Terms and Average Hypothesis Conditional Probability for GRMMAE with $\gamma = 1$



(b) Beta Terms and Average Hypothesis Conditional Probability for GRMMAE with $\gamma = 0$

Figure 3.3 Beta Terms and Average Hypothesis Conditional Probabilities for GRMMAEs with $\gamma = 1$ and $\gamma = 0$ and Correct Hypothesis Having Smallest Steady State Beta Term for $\gamma = 0$

3.8 Interpretation of β-stripping Using GRMMAE Derivation

As shown in Section 3.6, the GRMMAE design parameter, γ, affects only the $\beta_j^*(t_i)$ normalization terms. When β-dominance exists, it is common to use β-stripping to compensate for it, as described in Section 2.4.1. β-stripping is easily interpreted in the context of the GRMMAE derived in this chapter. Specifically, let γ vary among the elemental filters and over time. Under this assumption, each elemental filter is assigned an individual design parameter value $\gamma_j(t_i)$ at each measurement epoch t_i. Next, choose $\gamma_j(t_i)$ such that

$$\text{abs}\big(|\mathbf{T}_j(t_i)|\big) = \beta_j\left(t_i^-\right) \tag{3.42}$$

Equation (3.31) can be used to show

$$\beta_j^*(t_i) = \text{abs}\big(|\mathbf{T}_j(t_i)|\big)^{-1}\beta_j\left(t_i^-\right) \tag{3.43}$$

From Equation (3.43), if $\gamma_j(t_i)$ is chosen according to Equation (3.42), $\beta_j^*(t_i)$ will equal 1, and the $\beta_j^*(t_i)$ terms will be stripped from the hypothesis conditional probability calculation. Thus, in the context of the GRMMAE, the β-stripping method is interpreted as choosing $\gamma_j(t_i)$ such that Equation (3.42) is satisfied. Since, for linear elemental filters, $\beta_j^*(t_i)$ is pre-computable using the elemental filter models, calculating $\gamma_j(t_i)$ from Equation (3.42) is also pre-computable.

3.9 Chapter Summary

This chapter defined the generalized residual vector. It gave a derivation of the properties for the generalized residual vector for both the correct and incorrect elemental filters. It described the effect of the transformation matrix $\mathbf{T}(t_i)$. Next, it provided a derivation of the GRMMAE's hypothesis conditional probability formula. Then, the relationship between this generalized hypothesis conditional probability $p_k^*(t_i)$ and the traditional hypothesis conditional probability $p_k(t_i)$ was derived. Most importantly, it demonstrated that the effect of the γ design parameter is to increase or decrease the $\beta^*(t_i)$ term mag-

nitude artificially in the Gaussian densities used to calculate the hypothesis conditional probability. Because of this, the GRMMAE is susceptible to the β-dominance effect. Finally, this chapter concluded with a derivation of the generalized hypothesis conditional probability formula for applications in which the transformation matrix $\mathbf{T}(t_i)$ varies and does not vary among the elemental filters. In the case when $\mathbf{T}(t_i)$ does not vary among elemental filters, it demonstrated that the GRMMAE is equivalent to a traditional MMAE, regardless of the γ value. In the case when $\mathbf{T}(t_i)$ varies among the elemental filters, the generalized residual hypothesis conditional probability formula remains dependent upon γ making those applications susceptible to β-dominance if $\gamma \neq 1$. Since the GRMMAE is equivalent to a traditional MMAE for non-varying $\mathbf{T}(t_i)$ applications, and since the GRMMAE is susceptible to β-dominance in all other applications, it is not recommended for most applications. However, as derived in this chapter, the γ design parameter gives an ability to exert some measure of control over the magnitude of the Gaussian density $\beta_j^*(t_i)$ terms. Therefore, the GRMMAE may be useful in applications where β-dominance is expected. This is explored further in Sections 4.4 and 5.2. Finally, all sections in this chapter are research contributions of this dissertation.

IV. Simulations and Analysis

4.1 Residual Properties Verification

Sections 4.1.2 through 4.1.4 present examples to verify the generalized residual vector and generalized residual covariance matrix equations derived in Chapter III. These examples are presented in the same order as the derivations in Sections 3.3.1 and 3.3.4 with which they correspond. The examples all use a common truth model which is described in Section 4.1.1. Since GPS carrier-phase integer ambiguity resolution is the application that inspired the GRMMAE development, the truth model is designed to mimic that application. However, in an effort to keep the GRMMAE demonstration general, GPS carrier-phase integer ambiguity resolution is not specifically used in these sections. Ambiguity resolution will be used in Section 4.2 to demonstrate another GRMMAE property. In each of these examples, a different parameter is assumed to be unknown, and a GRMMAE is built to determine the correct parameter value given in the common truth model.

4.1.1 Truth Model. The truth model for the examples in Sections 4.1.2 through 4.1.4 is a simple 1-D system. It describes a vehicle moving along a line. The system state vector consists of position, velocity, and acceleration states, all along the line. The dynamics model is a first order Gauss-Markov acceleration (FOGMA) model given by

$$\dot{\mathbf{x}}(t) \;=\; \mathbf{F}(t)\,\mathbf{x}(t) \;+\; \mathbf{w}(t) \tag{4.1}$$

where $\mathbf{x}(t)$ is the state vector consisting of the position, velocity, and acceleration states

$$\mathbf{x}(t) \;=\; \begin{bmatrix} \mathsf{p}(t) \\ \mathsf{v}(t) \\ \mathsf{a}(t) \end{bmatrix} \tag{4.2}$$

and $\mathbf{F}(t)$ is the dynamics matrix given by

$$\mathbf{F}(t) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -1/t_a \end{bmatrix} \tag{4.3}$$

where t_a is the acceleration time constant. Finally, in Equation (4.1), $\mathbf{w}(t)$ is the continuous time dynamics noise given by

$$\mathbf{w}(t) = \begin{bmatrix} 0 \\ 0 \\ \mathsf{w_a}(\mathsf{t}) \end{bmatrix} \tag{4.4}$$

where $\mathsf{w_a}(\mathsf{t})$ is the zero-mean, white, Gaussian noise driving the acceleration state. The covariance kernel of the noise vector is given by

$$E\left[\mathbf{w}(t)\,\mathbf{w}^T(\tau)\right] = \mathbf{Q}(t)\,\delta(t - \tau) \tag{4.5}$$

The FOGMA model is commonly used in navigation applications including GPS carrier-phase integer ambiguity resolution. It assumes that the vehicle acceleration is well described with a first-order lag driven by white, Gaussian noise. From Equations (4.1) through (4.4), the model for the acceleration state is given by

$$\frac{d}{dt}\,\mathsf{a}(\mathrm{t}) = -\frac{1}{\mathrm{t_a}}\,\mathsf{a}(\mathrm{t}) + \mathsf{w_a}(\mathrm{t})$$

where $\mathsf{a}(\mathsf{t})$ is the vehicle acceleration in 1-D and $\mathsf{w_a}(\mathsf{t})$ is the continuous-time white, Gaussian noise driving the model. Changing the acceleration time constant and/or the strength, $\mathbf{Q}(t)$, of the driving noise changes the vehicle motion described by the truth model. A true value of $t_a = 2$ is arbitrarily chosen for these examples. The strength of $\mathsf{w_a}(\mathsf{t})$ is used in the calculation of the equivalent discrete-time noise covariance matrix $\mathbf{Q}_d(t_i)$, which will be discussed shortly.

Since the dynamics matrix in Equation (4.3) is time-invariant, the state transition matrix is calculated using the matrix exponential [40]

$$\mathbf{\Phi}\left(t_i, t_{i-1}\right) \;= e^{\mathbf{F}(t)\Delta t} \tag{4.6}$$

where $\Delta t = t_i - t_{i-1}$.

Using $t_a = 2$ and Equations (4.3) and (4.6), the true state dynamics model is

$$\mathbf{x}\left(t_{i+1}\right) \;=\; \begin{bmatrix} 1 & 1 & 0.426 \\ 0 & 1 & 0.787 \\ 0 & 0 & 0.607 \end{bmatrix} \mathbf{x}\left(t_i\right) \;+\; \mathbf{w}_d\left(t_i\right) \tag{4.7}$$

In Equation (4.7), $\mathbf{w}_d\left(t_i\right)$ is the equivalent discrete-time dynamics driving noise calculated from the continuous-time noise $\mathbf{w}\left(t\right)$. The equivalent discrete-time noise is zero-mean, white, Gaussian with covariance of $\mathbf{Q}_d\left(t_i\right)$ calculated from $\mathbf{Q}\left(t\right)$ [29]

$$\mathbf{Q}_d\left(t_i\right) \;=\; \int_{t_i}^{t_{i+1}} \mathbf{\Phi}\left(t_{i+1}, \tau\right) \mathbf{Q}\left(\tau\right) \mathbf{\Phi}^T\left(t_{i+1}, \tau\right) d\tau \tag{4.8}$$

Two independent measurement sensors are used for these examples. Each sensor measures the 1-D position of the vehicle along a line. The measurement model is given by

$$\mathbf{z}\left(t_i\right) \;=\; \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \mathbf{x}\left(t_i\right) \;+\; \mathbf{v}\left(t_i\right) \tag{4.9}$$

where $\mathbf{v}\left(t_i\right)$ is the measurement noise vector.

The equivalent discrete-time dynamics driving noise and measurement noise are independent, zero-mean, white, Gaussian noises with covariances of $\mathbf{Q}_d\left(t_i\right)$ and $\mathbf{R}\left(t_i\right)$, respectively. Two noise scenarios are used, a High $\mathbf{Q}_d$/Low $\mathbf{R}$ scenario (which is similar to

the GPS carrier-phase ambiguity resolution application) and a Low $\mathbf{Q}_d$/High $\mathbf{R}$ scenario. These relative values are chosen to create scenarios in which the elemental filters emphasize one model (dynamics or measurement) over the other. The noise covariance matrices for these scenarios are

$$\begin{array}{cc} \textbf{High } \mathbf{Q}_d\textbf{/Low R Scenario} & \textbf{Low } \mathbf{Q}_d\textbf{/High R Scenario} \\ \mathbf{Q}_d(t_i) = \begin{bmatrix} 0.0454 & 0.0838 & 0.0646 \\ 0.0838 & 0.1548 & 0.1193 \\ 0.0646 & 0.1193 & 0.0920 \end{bmatrix} & \mathbf{Q}_d(t_i) = \begin{bmatrix} 0.0454 & 0.0838 & 0.0646 \\ 0.0838 & 0.1548 & 0.1193 \\ 0.0646 & 0.1193 & 0.0920 \end{bmatrix} \times 10^{-4} \\ \mathbf{R}(t_i) = \begin{bmatrix} 0.0025 & 0 \\ 0 & 0.0025 \end{bmatrix} & \mathbf{R}(t_i) = \begin{bmatrix} 25 & 0 \\ 0 & 25 \end{bmatrix} \end{array} \tag{4.10}$$

The initial conditions for each scenario are

$$\mathbf{x}_0 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{P}_0 = \begin{bmatrix} 25 & 0 & 0 \\ 0 & 100 & 0 \\ 0 & 0 & 10 \end{bmatrix} \tag{4.11}$$

Finally, in each example that follows (including the measurement bias example) the true measurements are unbiased. Thus, in the measurement bias example, the true bias is given by

$$\mathbf{n}_{true} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{4.12}$$

The truth model given in Equations (4.1) through (4.12) is used for the examples in the next four sections. These examples demonstrate the generalized residual mean and covariance equations derived in Sections 3.3.1 through 3.3.4. The examples are presented in the same order as the derivations in Chapter III. The GRMMAEs applied to this truth model vary according to the unknown parameter being estimated. The GRMMAE models for each example are given in the corresponding section.

4.1.2 Example: Unknown Measurement Bias Scenario. This example is used to demonstrate the residual properties derived in Section 3.3.1. The truth model presented in Section 4.1.1 is used for this example. The elemental filters in the GRMMAE are identical but act upon corrected measurements given by Equation (3.7), repeated here for convenience

$$\tilde{\mathbf{z}}_j(t_i) = \mathbf{z}(t_i) - \mathbf{n}_j \tag{3.7}$$

where $\mathbf{z}(t_i)$ is the biased sensor measurement, and $\mathbf{n}_j$ is the hypothesized bias for the j^{th} elemental filter. Thus, the elemental filter model for this example is given by Equations 4.1 through 4.12. Three elemental filters are used, hypothesizing different measurement biases, which is very similar to the carrier-phase integer ambiguity resolution approach in [17, 18]. The hypothesized measurement biases are

$$\mathbf{n}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \mathbf{n}_2 = \begin{bmatrix} 5 \\ 10 \end{bmatrix} \qquad \mathbf{n}_3 = \begin{bmatrix} 0 \\ 5 \end{bmatrix} \tag{4.13}$$

where n_j is the hypothesized bias vector implemented in the j^{th} elemental filter ($j = 1, 2, 3$). The true measurement bias is $\mathbf{n}_1$.

This system was simulated 1000 times, running each simulation for 100 measurement epochs. Sample averages of the generalized residual vectors and the generalized residual covariance matrices were calculated at each measurement epoch. Thus, the 1000 Monte Carlo runs were used to yield a 100-step time history of 2-dimensional mean vectors (2 measurement sources, 100 measurement epochs) and 100 2 x 2 sample covariance matrices (one for each epoch). Additionally, sample averages of the state vectors prior to measurement incorporation were calculated for each elemental filter. Thus, 100 state vectors (one for each epoch), $\widehat{\mathbf{x}}_j\left(t_i^-\right)$, were produced for each of the 3 elemental filters. Since the $\mathbf{H}(t_i)$ and $\mathbf{K}(t_i)$ matrices do not vary from one Monte Carlo run to the next, it is not necessary to calculate sample averages of $\mathbf{H}(t_i)$, $\mathbf{K}(t_i)$, or $\mathbf{T}(t_i)$. Additionally, since the hypothesized measurement biases do not change from one Monte Carlo run to the next, it is not necessary to compute sample averages of the difference in the hypothesized biases

$\boldsymbol{\Delta}\mathbf{n}_{kj}$. This process was repeated a total of 4 times, once for each combination of the noise covariance matrices in Equation (4.10) with γ equal to both 0 and 1. This is depicted graphically by the "X" marks in Table 4.1

For each Monte Carlo simulation listed in Table 4.1, the sample averages were used to calculate the mean of the generalized residual vector for the incorrect elemental filters (2 and 3) using Equation (3.11), a modified portion of which is repeated here for convenience

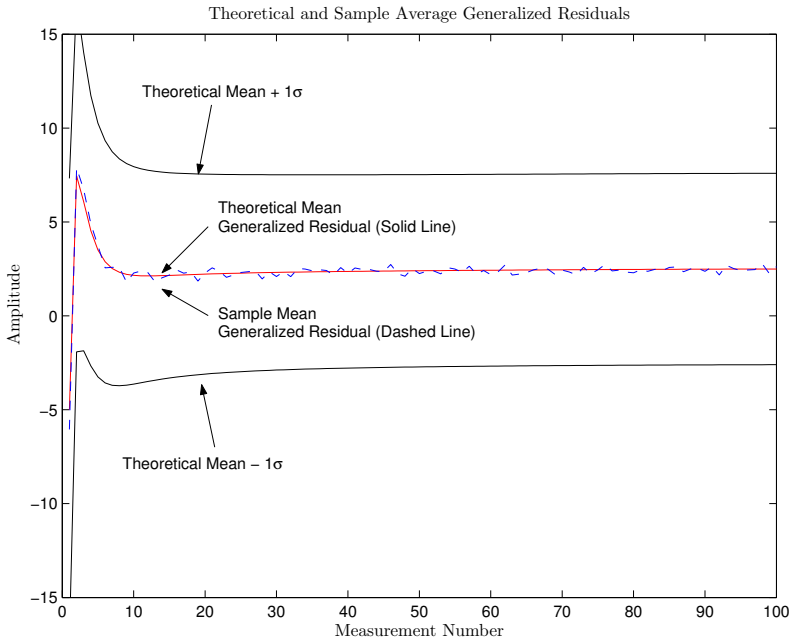
$$\mathbf{m}_{\mathbf{r}^*_{j_{theory}}}(t_i) = \mathbf{T}(t_i)\left[\mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_{k_{aver}}\left(t_i^-\right) - \widehat{\mathbf{x}}_{j_{aver}}\left(t_i^-\right)\right) + \boldsymbol{\Delta}\mathbf{n}_{kj}\right] \tag{3.11}$$

where $\widehat{\mathbf{x}}_{k_{aver}}\left(t_i^-\right)$, and $\widehat{\mathbf{x}}_{j_{aver}}\left(t_i^-\right)$ are the sample averages taken from the 1000 Monte Carlo runs, and the subscripts *aver* and *theory* denote sample averages and a residual mean vector calculated using the theory from Chapter III, respectively. It was necessary to calculate sample averages of the state vector estimates, $\widehat{\mathbf{x}}_{k_{aver}}\left(t_i^-\right)$ and $\widehat{\mathbf{x}}_{j_{aver}}\left(t_i^-\right)$, because the true state varied in each Monte Carlo run due to the dynamics model noise. The subscripts *aver* and *theory* were added here to distinguish between realized random variable values in the 1000 Monte Carlo simulations (*aver*) and values calculated using the equations in Chapter III (*theory*).

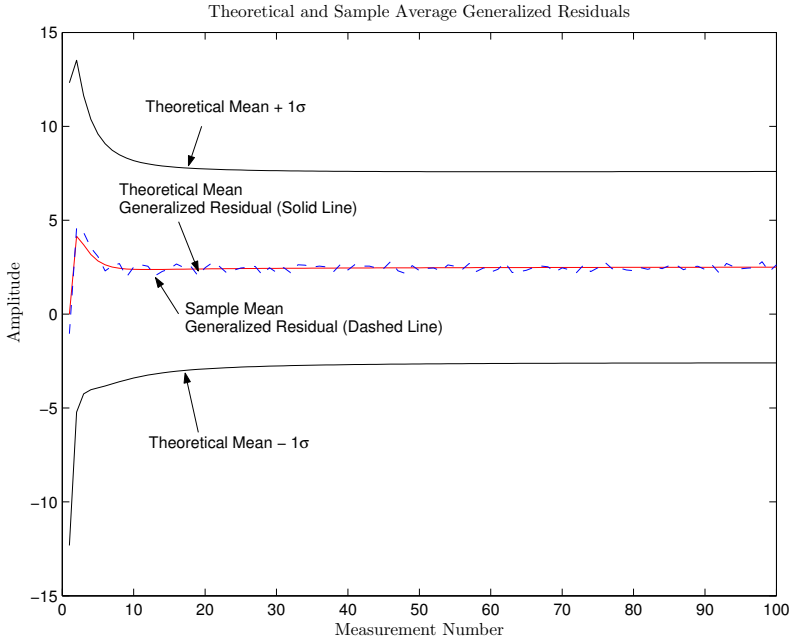
Figure 4.1 shows a plot of the theoretical generalized residuals, calculated using the modified version of Equation (3.11), and the realized sample average residual vectors for elemental filters 2 and 3 (the incorrect filters). In this figure, the Low $\mathbf{Q}_d$/High $\mathbf{R}$ noise scenario was used. For that scenario, the elemental Kalman filters place increased emphasis on the filter dynamics model and less emphasis on the measurements. Additionally, the traditional residual ($\gamma = 1$) was used. Finally, in order to make the figure more clear, only

Table 4.1 Matrix of Monte Carlo Test Scenarios: Combinations of Noise Matrices and γ Values

		Generalized Residual Type	
		Traditional ($\gamma = 1$)	**Post-Fit** ($\gamma = 0$)
Noise Scenario	**High $\mathbf{Q}_d$/Low R**	X	X
	Low $\mathbf{Q}_d$/High R	X	X



(a) Elemental Filter 2



(b) Elemental Filter 3

Figure 4.1 Theoretical and Sample Average Generalized Residuals from 1000 Simulations for Measurement Source 1 in Low $\mathbf{Q}_d(t_i)$/High $\mathbf{R}$ Noise Scenario with Traditional Residual ($\gamma = 1$) – Measurement Bias Example

the residuals associated with measurement source 1 are shown. The plot for measurement source 2 was similar. Figure 4.1 demonstrates that the sample mean generalized residuals track the theoretical mean generalized residuals from Equation (3.11). Relative to the generalized residual covariance, the realized sample mean closely matches the theoretical mean. The simulations for the other combinations of noise covariance matrices and gamma values in Table 4.1 produced results similar to Figure 4.1 and are not shown. Figure 4.1 demonstrates that Equation (3.11) correctly calculated the mean value for the generalized residuals from the incorrect elemental filters, based on the traditional residual from the correct elemental filter.

Next, the covariance of the generalized residuals from the incorrect elemental filters was confirmed. For the measurement bias application, this formula is given by Equation (3.13), a modified version of which is repeated here for convenience

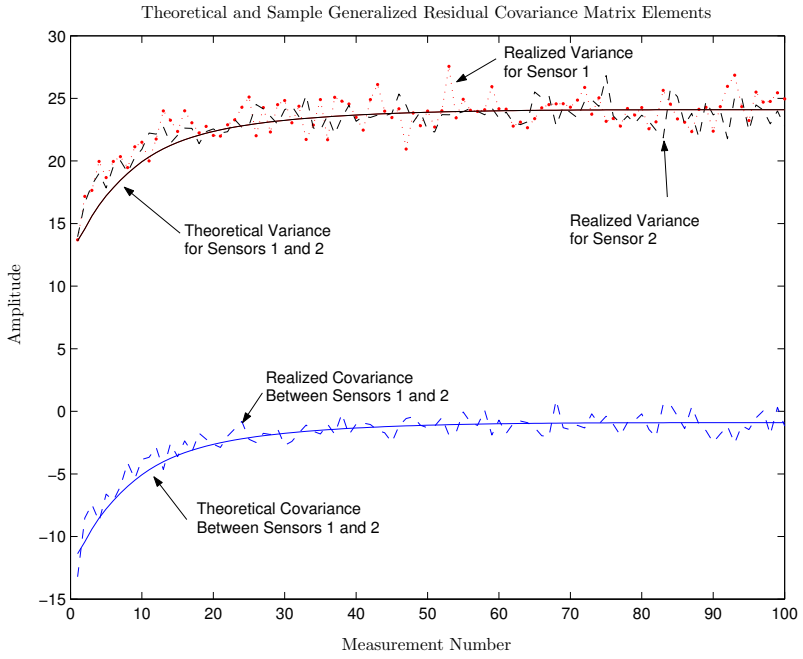
$$\mathbf{A}^*_{j_{theory}}(t_i) = \mathbf{T}(t_i)\,\mathbf{A}_{k_{filt}}\left(t_i^-\right)\mathbf{T}^T(t_i) \tag{3.13}$$

where the subscript *theory* indicates a matrix calculated using the theory in Chapter III. The new subscript $filt$ indicates that the elemental filter-computed covariance was used. The filter-computed residual covariance was given immediately after Equation (2.47) and is

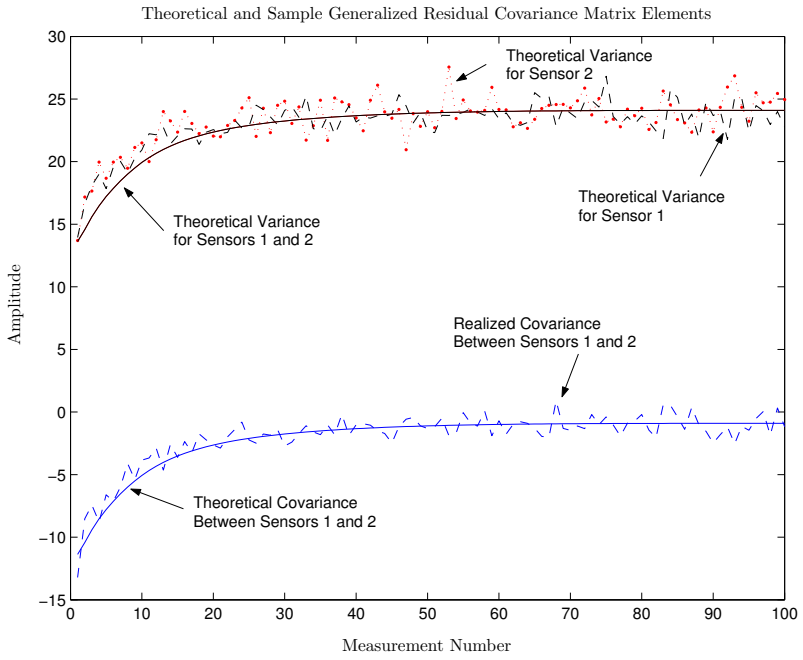
$$\mathbf{A}_{k_{filt}}\left(t_i^-\right) = \mathbf{H}(t_i)\,\mathbf{P}_k\left(t_i^-\right)\mathbf{H}^T(t_i) + \mathbf{R}(t_i) \tag{4.14}$$

None of the matrices above vary from one Monte Carlo run to the next. Therefore, there is no need to calculate a sample average value for the elemental filter-computed covariance matrix.

Figure 4.2 shows a representative plot of the realized sample covariance matrix and the theoretical covariance matrix calculated from Equation (3.13). Since there were two measurement sources, the residual covariance matrices were 2 x 2. In Figure 4.2, the line marked "Theoretical Variance for Sensors 1 and 2" is a plot of the variance values from the diagonal of the covariance matrix. In this case, the two variance values are the



(a) Elemental Filter 2



(b) Elemental Filter 3

Figure 4.2 Theoretical and Sample Average Generalized Residual Covariance Matrix Elements from 1000 Simulations in Low $\mathbf{Q}_d$/High $\mathbf{R}$ Noise Scenario with Post-Fit Residuals ($\gamma = 0$) – Measurement Bias Example

same and only one line appears in the plot, representing both sensors. The line marked "Theoretical Covariance for Sensors 1 and 2" is a plot of the off-diagonal elements of the covariance matrix. Since covariance matrices are symmetric, there is only a need to plot a single line for the covariance values. The plots of "Realized Variance" and "Realized Covariance" are the values from the realized sample covariance matrices corresponding to the theoretical values. Figure 4.2 shows the variance/covariance values for the Low $\mathbf{Q}_d$/High $\mathbf{R}$ noise scenario using post-fit residuals ($\gamma = 0$). Figure 4.2 demonstrates that the realized sample variance and covariance values track the corresponding theoretical values with errors that are an order of magnitude smaller than the theoretical values. This behavior was true for all combinations of noise covariance matrices and gamma values in Table 4.1, and indicates that Equation (3.13) correctly calculated the generalized residual covariance matrix. However, Figure 4.2 is shown because it demonstrates unexpected behavior. Typically, the residual variance is large at first and settles down to a smaller steady state value. In Figure 4.2, though, the residual variance values start smaller and increase to their steady state value. This was caused by the transformation matrix $\mathbf{T}(t_i)$. This unexpected behavior will be discussed in Chapter V in relation to a recommendation for future research into use of the GRMMAE for β-dominance compensation.

Finally, based on Equations (3.10) and (3.13), for this application, the generalized residuals from the incorrect elemental filters equal the generalized residuals from the correct elemental filter, offset by a bias. Figure 4.3 demonstrates this property. In this figure, the generalized residual associated with measurement source 1 from all three elemental filters is shown. Unlike previous figures, a single Monte Carlo run was used for this demonstration. This ensures that a direct comparison can be made between the elemental filters. If a sample average had been used, it is possible (though not likely) to have multiple runs which *individually* exhibit incorrect behavior while still maintaining correct *average* behavior. By using a single run, this possibility was avoided. Figure 4.3 shows the generalized residuals for the High $\mathbf{Q}_d$/Low $\mathbf{R}$ noise covariance scenario with post-fit residuals (i.e., $\gamma = 0$). In Figure 4.3, the generalized residuals from elemental filters 2 and 3 overlap due to the choice of measurement bias hypotheses, and thus, only one line appears in the plot representing both filters. A comparison of the residuals from elemental filter 1 (the correct hypothesis)

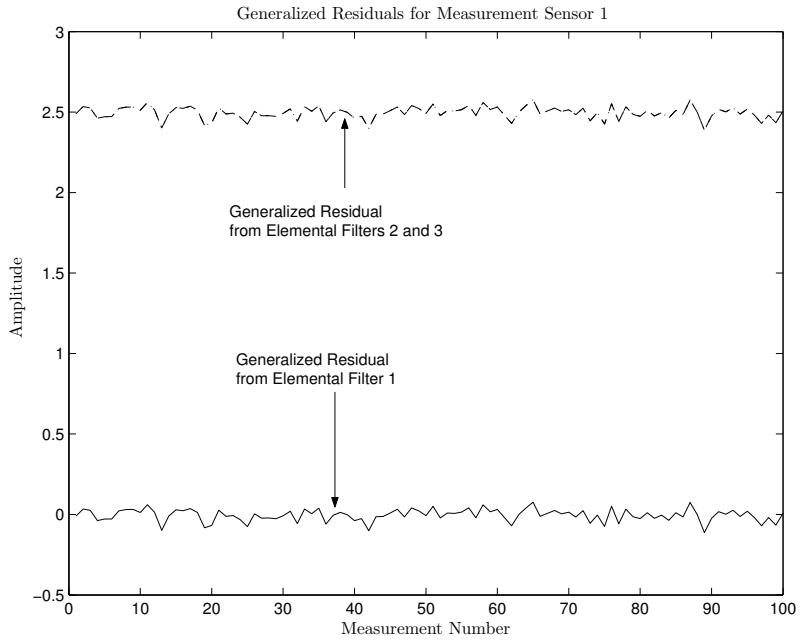


Figure 4.3 Single Run Generalized Residual for Measurement Sensor 1 in High $\mathbf{Q}_d$/Low $\mathbf{R}$ Noise Scenario with Post-Fit Residuals ($\gamma = 0$)

with those from elemental filters 2 and 3 clearly shows that the residuals from the incorrect elemental filters equal the residuals from the correct elemental filter, with an added bias.

4.1.3 Example: Unknown $\mathbf{Q}_d$ Scenario. This example demonstrates an application with an unknown equivalent discrete-time dynamics noise covariance matrix. Equations to calculate the generalized residuals for the incorrect elemental filters from the traditional residuals for the correct filter are given in Section 3.3.2. As with other examples, the truth model presented in Section 4.1.1 was used for this example.

An example such as this might arise, for example, in a target tracking problem. In this application, the FOGMA model in Equations (4.1) through (4.7) adequately describes the dynamics of a vehicle. However, changes in target dynamics are described by varying the strength of the continuous-time dynamics noise $w(t)$. For instance, a target at rest would be represented with a very small dynamics model noise strength. Various dynamic operating modes (accelerating, constant velocity, etc.) would be represented with other dynamics noise strengths. Those varying values of strength for $\mathbf{w}(t)$ lead to varying $\mathbf{Q}_d(t_i)$

for each operating mode. The GRMMAE is then used to determine which hypothesized $\mathbf{Q}_d(t_i)$ best matches the motion of the vehicle, allowing vehicle tracking.

In this example, three elemental filters were used, hypothesizing three values for the equivalent discrete-time noise covariance. The hypothesized values for $\mathbf{Q}_d(t_i)$ were arbitrarily chosen to be

$$\mathbf{Q}_d(t_i) \qquad 2.5\mathbf{Q}_d(t_i) \qquad 6.7\mathbf{Q}_d(t_i)$$

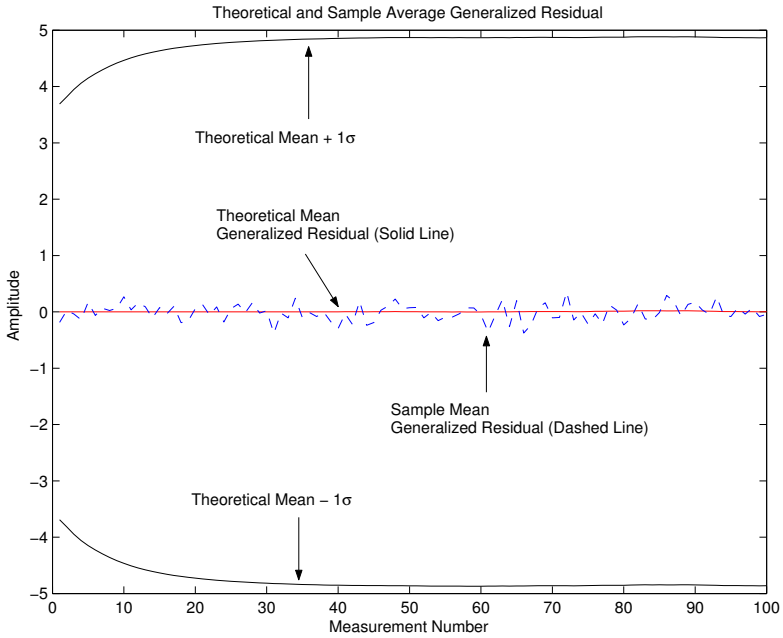
where $\mathbf{Q}_d(t_i)$ is the value given in Equation (4.10). Two noise scenarios and two values of γ were used as shown in Table 4.1, for a total of four sets of Monte Carlo simulations. Each scenario defined in Table 4.1 was simulated 1000 times with 100 measurement epochs used in each Monte Carlo simulation. As in the measurement bias example, sample averages of the variables necessary to confirm the Equations in Section 3.3.2 were calculated. Since the $\mathbf{H}(t_i)$ and $\mathbf{K}_j(t_i)$ matrices do not vary from one Monte Carlo run to the next, it was not necessary to calculate sample averages of $\mathbf{H}(t_i)$, $\mathbf{K}_j(t_i)$, or $\mathbf{T}_j(t_i)$. Notice, however, that unlike the previous example, $\mathbf{K}_j(t_i)$ and $\mathbf{T}_j(t_i)$ do vary among the elemental filters.

For each Monte Carlo simulation listed in Table 4.1, the sample averages were used to calculate the mean of the generalized residual vector for the incorrect elemental filters (2 and 3) using Equation (3.15), a modified portion of which is repeated here for convenience

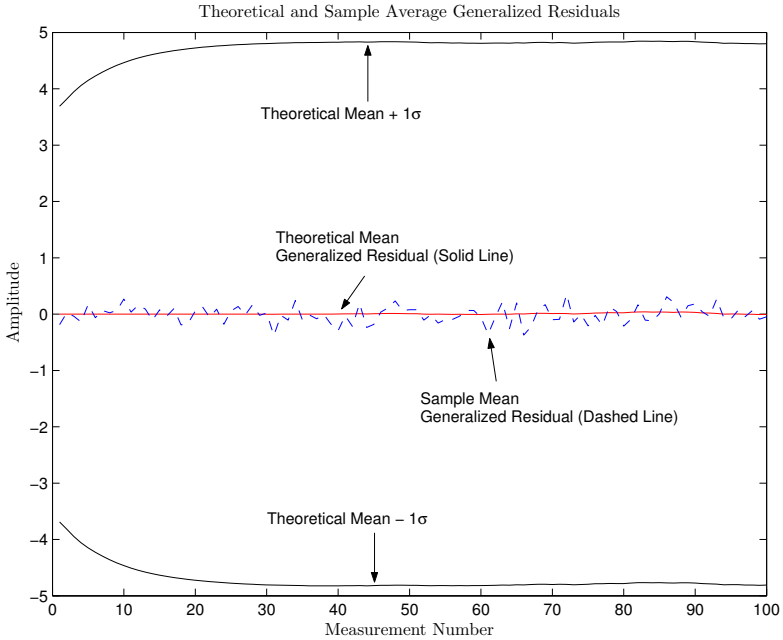
$$\mathbf{m}_{\mathbf{r}^*_{j_{theory}}}(t_i) = \mathbf{T}_j(t_i)\,\mathbf{H}(t_i)\left[\widehat{\mathbf{x}}_{k_{aver}}\left(t_i^-\right) - \widehat{\mathbf{x}}_{j_{aver}}\left(t_i^-\right)\right] \tag{3.15}$$

where $\widehat{\mathbf{x}}_{k_{aver}}\left(t_i^-\right)$, and $\widehat{\mathbf{x}}_{j_{aver}}\left(t_i^-\right)$ are the sample averages taken from the 1000 Monte Carlo runs, and the subscripts *aver* and *theory* denote sample averages and a residual mean vector calculated from the theory in Chapter III, respectively.

Figure 4.4 shows a plot of the theoretical generalized residuals, calculated using the modified version of Equation (3.15), and the realized sample average residual vectors for elemental filters 2 and 3 (the incorrect filters). In this figure, the Low $\mathbf{Q}_d$/High $\mathbf{R}$ noise scenario was used. For that scenario, the elemental Kalman filters place increased emphasis on the filter dynamics model and less emphasis on the measurements. Additionally,



(a) Elemental Filter 2



(b) Elemental Filter 3

Figure 4.4 Theoretical and Sample Average Generalized Residuals from 1000 Simulations for Measurement Source 1 in Low $\mathbf{Q}_d$/High $\mathbf{R}$ Noise Scenario with Post-Fit Residual ($\gamma = 0$) – Unknown $\mathbf{Q}_d$ Example

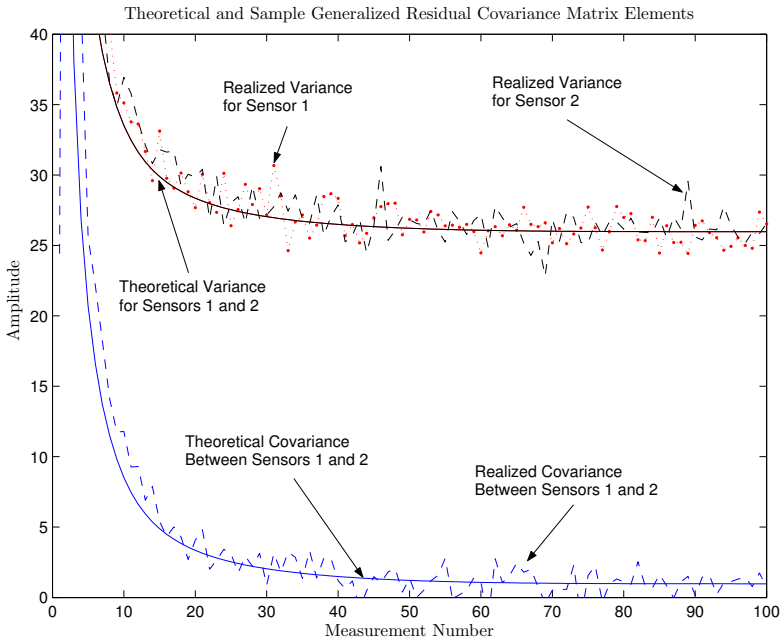
the post-fit residual ($\gamma = 0$) was used. Finally, in order to make the figure more clear, only the residuals associated with measurement source 1 are shown. Figure 4.4 shows that, relative to the generalized residual standard deviation, the realized sample mean generalized residual exhibited a good match to the theoretical mean. This demonstrates that Equation (3.15) correctly calculated the mean generalized residual vector for the incorrect elemental filters based on the traditional residual vector from the correct elemental filter.

Next, the covariance of the generalized residuals from the incorrect elemental filters was confirmed. For the unknown $\mathbf{Q}_d$ application, this formula is also given by Equation (3.15), a modified version of which is repeated here for convenience

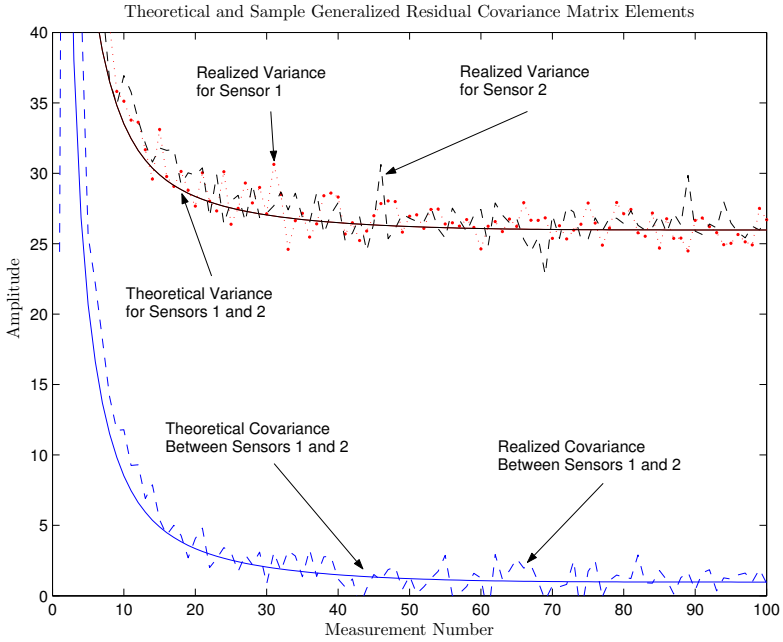
$$\mathbf{A}^*_{j_{theory}}(t_i) = \mathbf{T}_j(t_i)\,\mathbf{A}_{k_{filt}}\left(t_i^-\right)\mathbf{T}_j^T(t_i) \tag{3.15}$$

where the subscripts *theory* and *filt* indicate a covariance matrix calculated from the theory in Chapter III and the elemental filter-computed covariance matrix, respectively. As before, the filter-computed covariance matrix does not change from one run to the next. Therefore, there is no need to calculate a sample average value for it.

Figure 4.5 shows a representative plot of the realized sample covariance matrix and the covariance matrix calculated from Equation (3.15). In Figure 4.5, the line marked "Theoretical Variance for Sensors 1 and 2" is a plot of the variance values from the diagonal of the covariance matrix. In this case, the two variance values were the same and only 1 line appears in the plot, representing both sensors. The line marked "Theoretical Covariance for Sensors 1 and 2" is a plot of the off-diagonal elements of the covariance matrix. The plots of "Realized Variance" and "Realized Covariance" are the values from the realized sample covariance matrices corresponding to the theoretical values. Figure 4.5 shows the variance/covariance values for the Low $\mathbf{Q}_d$/High $\mathbf{R}$ noise scenario using traditional residuals ($\gamma = 1$). Figure 4.5 demonstrates that the realized sample variance and covariance values track the corresponding theoretical values with errors that are an order of magnitude smaller than the theoretical values. This behavior was true for all combinations of noise covariance matrices and gamma values in Table 4.1, and indicates that Equation (3.15) correctly calculated the generalized residual covariance matrix.



(a) Elemental Filter 2



(b) Elemental Filter 3

Figure 4.5 Theoretical and Sample Generalized Residual Covariance Matrix Elements from 1000 Simulations in Low $\mathbf{Q}_d$/High $\mathbf{R}$ Noise Scenario with Traditional Residuals ($\gamma = 1$) – Unknown $\mathbf{Q}_d$ Example

4.1.4 Example: Unknown $\mathbf{F}(t)$ *Scenario.* This example confirms the equations for the generalized residual vector and covariance matrix, derived in Section 3.3.3, for an application in which the matrix $\mathbf{F}(t)$ describing the homogeneous system dynamics is subject to uncertainty. This application could arise in a target identification problem. In that application, various targets would be described by the general FOGMA model given in Equation (4.3). Each hypothesized target would have a different acceleration time constant t_a, which in turn leads to each hypothesis having a different $\mathbf{F}(t)$ matrix. A target being tracked would then be identified by matching the observed motion to the best hypothesis for $\mathbf{F}(t)$.

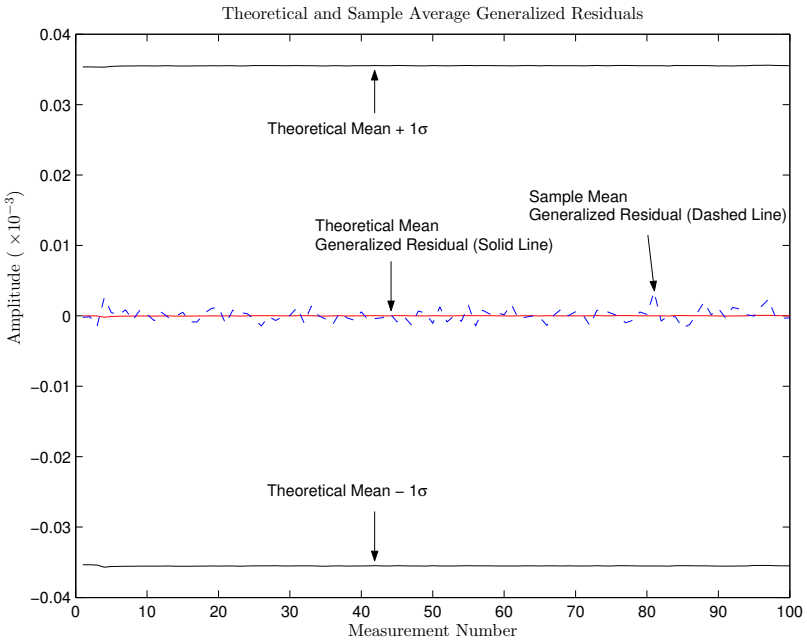
In this example, three elemental filters were used, hypothesizing three values for the acceleration time constant t_a. The hypothesized values for t_a were arbitrarily chosen to be 2, 4, and 8. Substitution of these acceleration constants into Equation (4.3), yields the three hypotheses for $\mathbf{F}(t)$

$$\mathbf{F}_1(t) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -0.5 \end{bmatrix} \quad \mathbf{F}_2(t) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -0.25 \end{bmatrix} \quad \mathbf{F}_3(t) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -0.125 \end{bmatrix} \tag{4.15}$$

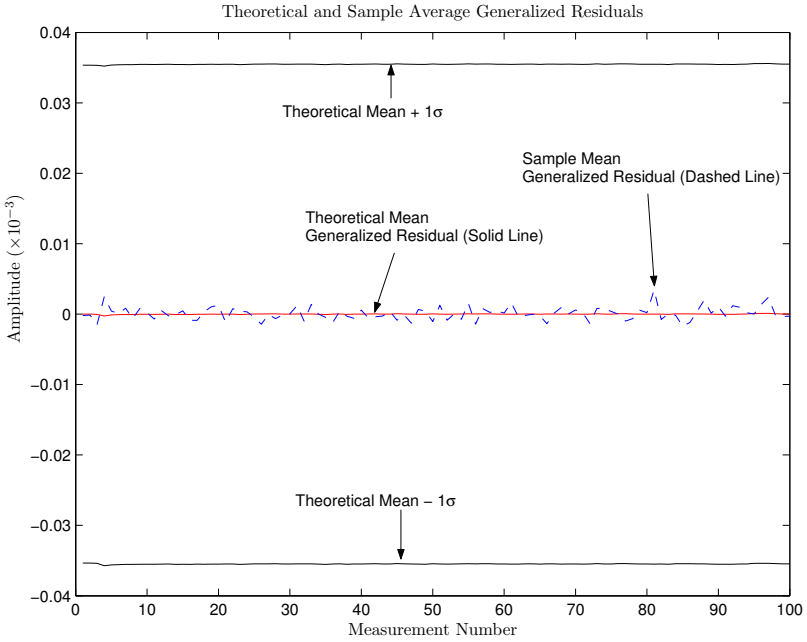
From Section 4.1.1, $\mathbf{F}_1(t)$ was the true value of the homogeneous dynamics matrix.

The analysis steps for the 1000 Monte Carlo simulations in Section 4.1.3 were repeated for this application, with similar results and conclusions. Figures 4.6 and 4.7 demonstrate that Equation (3.15) correctly calculated the mean vectors and covariance matrices of the generalized residuals from the incorrect elemental filters. Each figure is labelled according to the noise covariance scenario and type of generalized residual used.

4.1.5 Example: Unknown $\mathbf{R}$ *Scenario.* This example demonstrates an application with an unknown measurement noise covariance matrix. The equations for the generalized residual and its covariance for such an application are discussed in Section 3.3.4. This

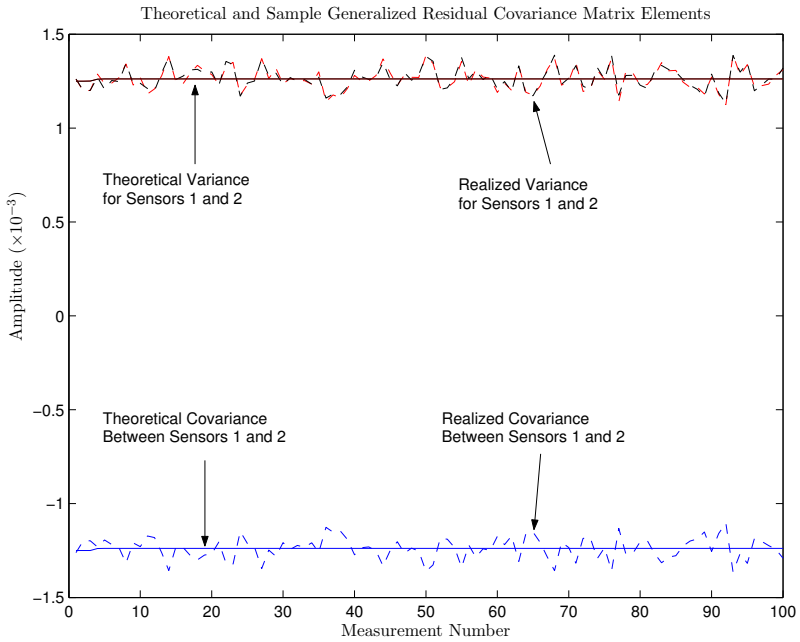


(a) Elemental Filter 2

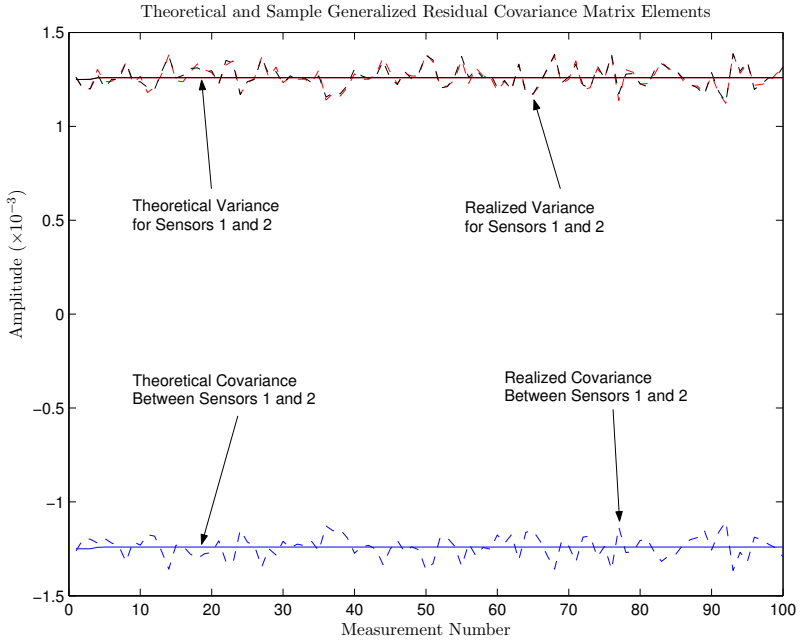


(b) Elemental Filter 3

Figure 4.6 Theoretical and Sample Average Generalized Residuals from 1000 Simulations for Measurement Source 1 in High $\mathbf{Q}_d$/Low $\mathbf{R}$ Noise Scenario with Post-Fit Residual ($\gamma = 0$) – Unknown $\mathbf{F}(t)$ Example



(a) Elemental Filter 2



(b) Elemental Filter 3

Figure 4.7 Theoretical and Sample Generalized Residual Covariance Matrix Elements from 1000 Simulations in High $\mathbf{Q}_d$/Low $\mathbf{R}$ Noise Scenario with Post-Fit Residuals ($\gamma = 0$) – Uknown $\mathbf{F}(t)$ Example

section presents results to confirm the validity of those equations. As with other examples, the truth model presented in Section 4.1.1 was used for this example.

An example of such an application from the navigation field is navigation using a GPS aided Inertial Navigation System (INS) in a GPS jamming environment. As jamming strength increases, the accuracy of the GPS receiver decreases. This is modelled by increasing the measurement noise covariance $\mathbf{R}(t_i)$. If the jamming becomes strong enough, GPS signal tracking is lost altogether, and the INS must run without GPS updates, which is modelled by another measurement noise covariance level. A GRMMAE that can determine the appropriate measurement noise covariance matrix would allow the navigation system to continue operation through a degradation and eventual loss of the GPS sensor. Additionally, it would automatically account for the re-introduction of GPS as the system left the jamming field.

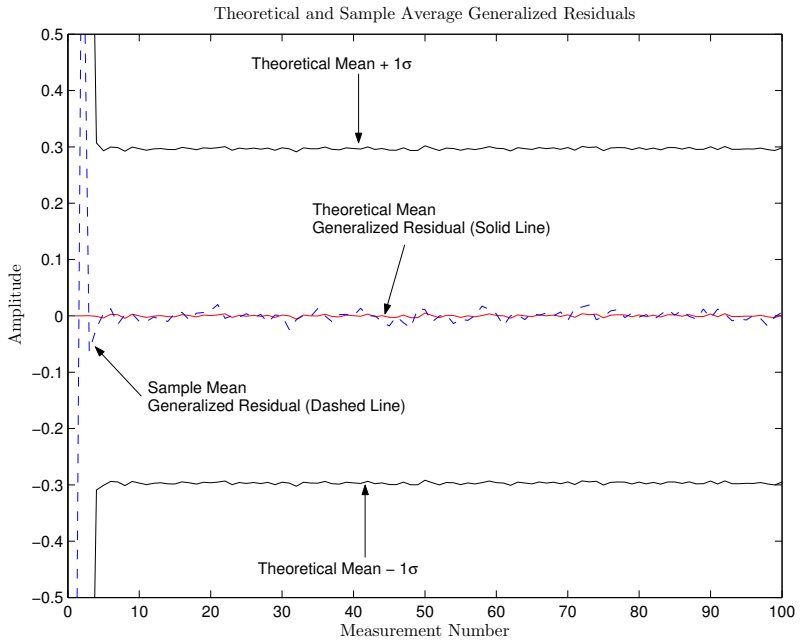
In this example, three elemental filters were used, hypothesizing three values for the measurement noise covariance matrix. The hypothesized values for $\mathbf{R}$ were arbitrarily chosen to be

$$\mathbf{R}(t_i) \qquad 2.5\mathbf{R}(t_i) \qquad 6.7\mathbf{R}(t_i)$$

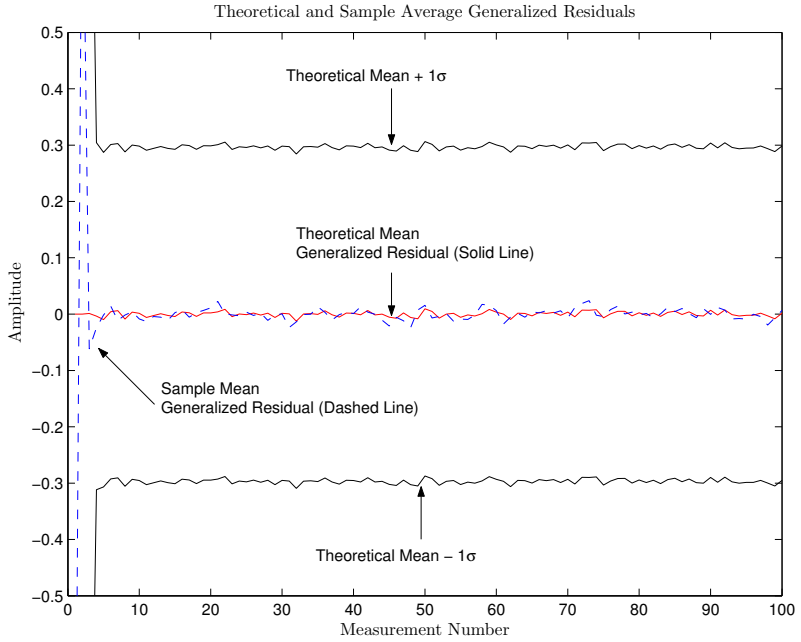
where $\mathbf{R}(t_i)$ is the value given in Equation (4.10).

The analysis steps for the 1000 Monte Carlo simulations in Section 4.1.3 were repeated for this application, with similar results and conclusions. Figures 4.8 and 4.9 demonstrate that Equation (3.15) correctly calculated the mean vectors and covariance matrices of the generalized residuals from the incorrect elemental filters. Each figure is labelled according to the noise covariance scenario and type of generalized residual used.

4.1.6 Measurement Bias Parameter with $\gamma = 0.5$. In the previous four examples, the GRMMAE design parameter γ was always set to 0 or 1 (post-fit or traditional residuals, respectively). This was done arbitrarily because these two values make the most physical sense. However, at no point in the GRMMAE derivation was γ constrained to be 0 or 1. In fact, γ is allowed to be any real, scalar value. Therefore, the measurement bias example was repeated for the High $\mathbf{Q}_d$/Low $\mathbf{R}$ noise scenario of Equation (4.10) with $\gamma = 0.5$

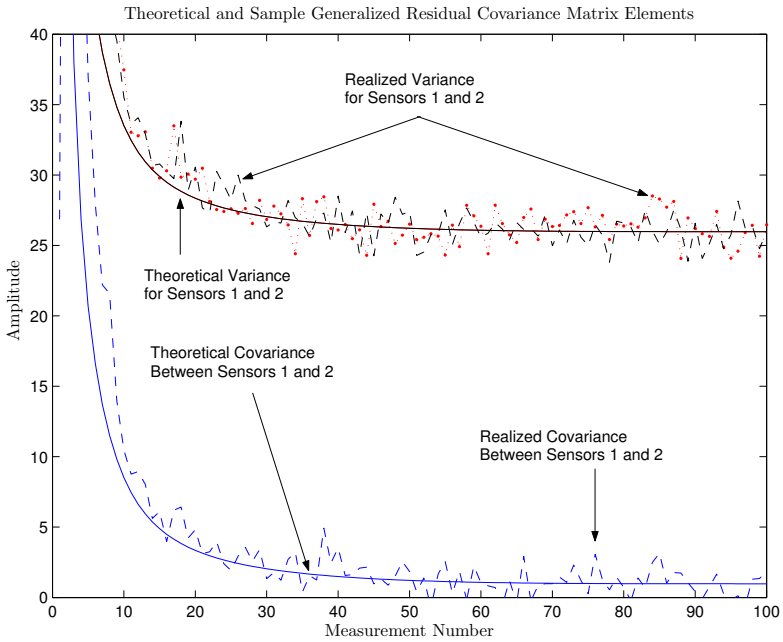


(a) Elemental Filter 2

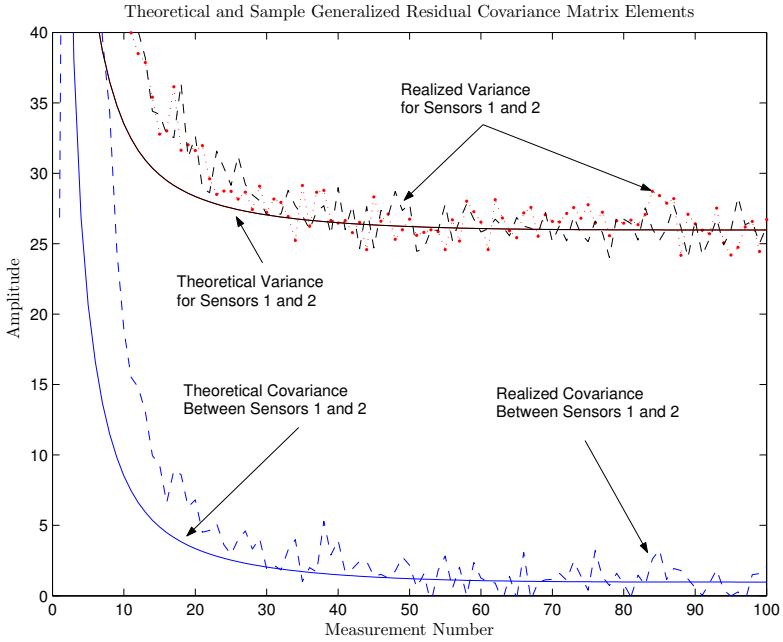


(b) Elemental Filter 3

Figure 4.8 Theoretical and Sample Average Generalized Residuals from 1000 Simulations for Measurement Source 1 in High $\mathbf{Q}_d$/Low $\mathbf{R}$ Noise Scenario with Traditional Residual ($\gamma = 1$) – Unknown $\mathbf{R}$ Example



(a) Elemental Filter 2



(b) Elemental Filter 3

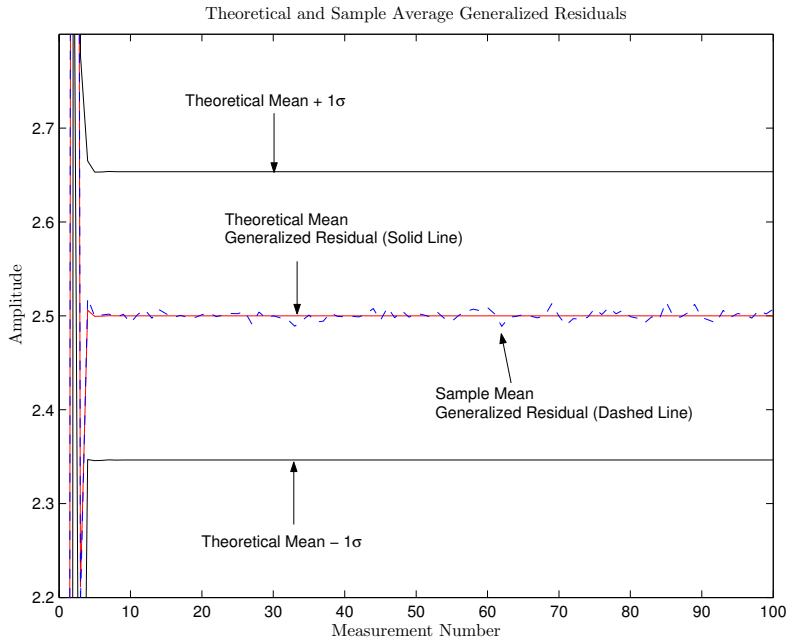
Figure 4.9 Theoretical and Sample Generalized Residual Covariance Matrix Elements from 1000 Simulations in Low $\mathbf{Q}_d$/High $\mathbf{R}$ Noise Scenario with Traditional Residuals ($\gamma = 1$) – Unknown $\mathbf{R}$ Example

(i.e., a linear combination of the post-fit and traditional residuals). The results of these 1000 Monte Carlo runs, shown in Figures 4.10 and 4.11, were similar to those in previous sections, demonstrating that Equations (3.11) and (3.13) for the mean and covariance of $\mathbf{r}_j^*(t_i)$ hold for $\gamma = 0.5$, as expected.

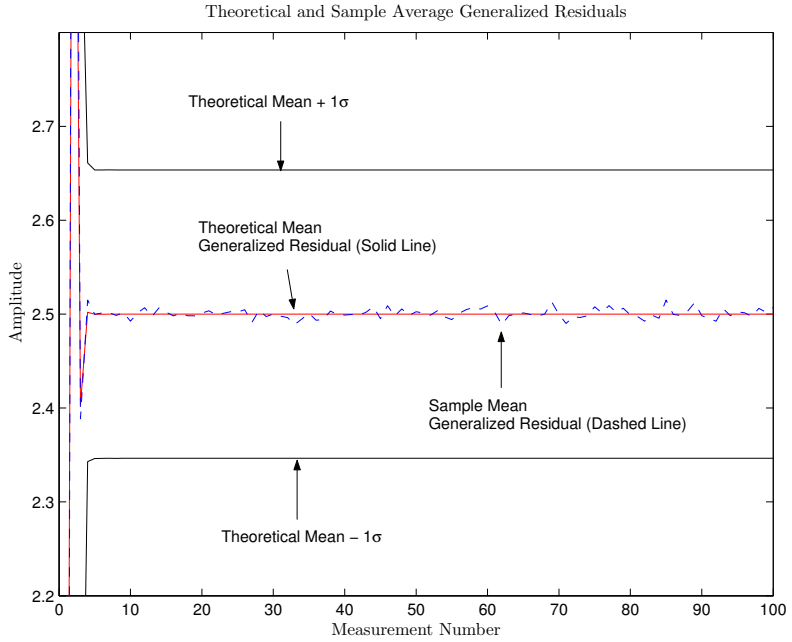
4.2 GRMMAE for GPS Carrier-Phase Integer Ambiguity Resolution

Carrier-phase integer ambiguity resolution is an example of a GRMMAE application with a constant measurement bias parameter. This general class of problems is presented in Sections 3.3.1 and 3.7.1. Previous research demonstrated that a GRMMAE with $\gamma = 0$ (i.e., using post-fit residuals) will successfully resolve GPS carrier-phase integer ambiguities [17, 18]. A research contribution of this dissertation is the derivation in Section 3.7.1 which shows that the hypothesis conditional probabilities calculated for this application are independent of the choice of γ. Thus, while the use of post-fit residuals is valid for this application, the use of traditional residuals is equally valid. This section will use GPS carrier-phase measurement data to demonstrate that GRMMAE's with $\gamma = 0$ and $\gamma = 1$ produce equivalent results when the parameter of interest is an unknown measurement bias.

Before proceeding, it is important to recognize that the GRMMAE for GPS carrier-phase integer ambiguity resolution does not meet the linearity assumption in Sections 3.3.1 and 3.7.1. In those sections, it is assumed that the GRMMAE elemental filters are linear Kalman filters. However, the measurement model for GPS is a nonlinear model, leading to the use of extended Kalman filters in the GRMMAE (recall that use of extended Kalman filters is a non-optimal, ad hoc extension to the MMAE and GRMMAE). Therefore, the measurement matrices $\mathbf{H}_j(t_i)$ in the elemental filters are first-order approximations as given in Equation (2.26) and Table 2.2. Because the measurement matrices are linearized around the current state estimate, and each elemental filter produces a unique position estimate, the measurement matrices in the elemental filters will not be equal in all elemental filters as assumed in Section 3.7.1. This, in turn, leads directly to the transformation matrices $\mathbf{T}_j(t_i)$ varying in the elemental filters, which violates the key assumption that

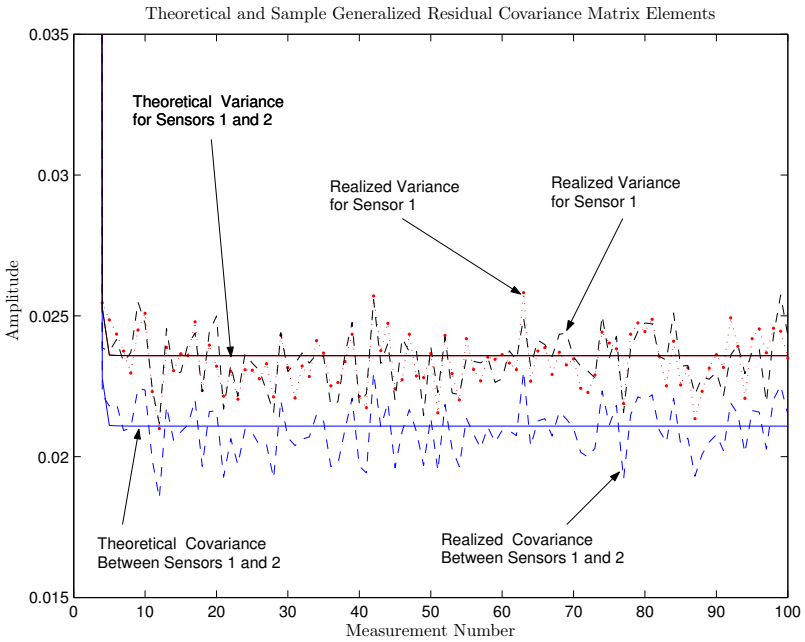


(a) Elemental Filter 2

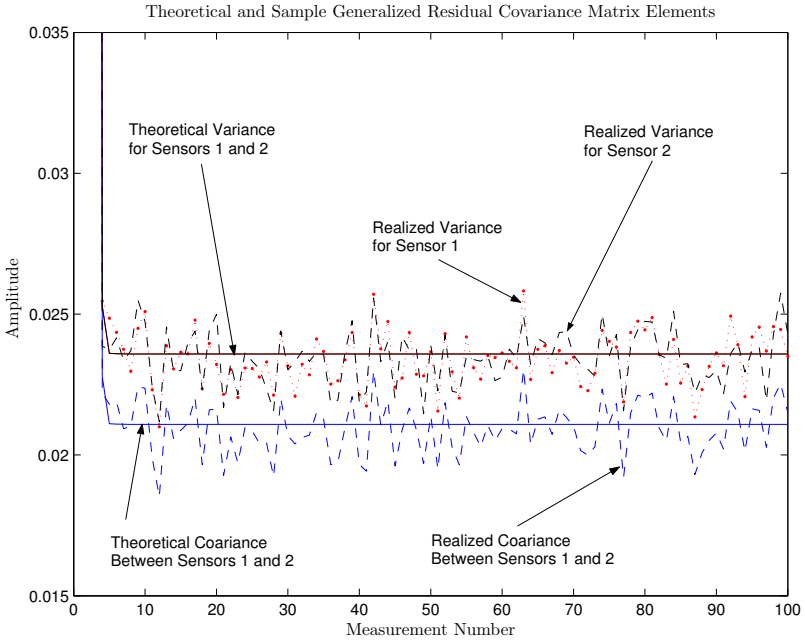


(b) Elemental Filter 3

Figure 4.10 Theoretical and Sample Average Generalized Residuals from 1000 Simulations for Measurement Source 1 in High $\mathbf{Q}_d$/Low $\mathbf{R}$ Noise Scenario with $\gamma = 0.5$ – Measurement Bias Example



(a) Elemental Filter 2



(b) Elemental Filter 3

Figure 4.11 Theoretical and Sample Generalized Residual Covariance Matrix Elements from 1000 Simulations in High $\mathbf{Q}_d$/Low $\mathbf{R}$ Noise Scenario with $\gamma = 0.5$ – Measurement Bias Example

$\mathbf{T}_j(t_i) = \mathbf{T}_k(t_i)$ for all j and k. However, because of measurement geometry, this effect is not significant for the GPS carrier-phase integer ambiguity resolution application.

When performing GPS carrier-phase integer ambiguity resolution, the measurement matrices are composed of unit line of sight vectors pointing from the estimated user position toward the satellites [39]. However, the user position estimates (produced using the hypothesized ambiguity vectors) will be separated by short distances on the order of centimeters. Since the GPS satellites are tens of millions of meters away from the user and the hypothesized user positions are separated by centimeters, the differences in the unit line of site vectors (and, hence the measurement matrices) are so small that they can be neglected. Thus, while the measurement matrices are not mathematically equal in all elemental filters, the differences are negligible and the results of Sections 3.3.1 and 3.7.1 apply to this application. While Section 3.7.1 shows that the transformation matrix $\mathbf{T}_j(t_i)$ can be eliminated from the hypothesis conditional probability calculation, it will be used in the results presented in this section. Thus, the effect of neglecting the differences in $\mathbf{T}_j(t_i)$ will be seen. The analysis could have been repeated implementing the assumption that $\mathbf{T}_j(t_i)$ does not vary among the elemental filters and removing it from hypothesis conditional probability calculation. However, the same result (in the next paragraph) would have been achieved. Therefore, $\mathbf{T}_j(t_i)$ for all elemental filters was used in the analysis that follows.

To demonstrate that any γ value results in equivalent hypothesis conditional probabilities, a test was performed. GPS double difference widelane carrier-phase measurements were used in GRMMAEs, with $\gamma = 0$ and $\gamma = 1$ (i.e., using post-fit and traditional residuals, respectively). The double difference measurements were taken over a 69 km baseline. In both cases, 100 elemental filters were built upon 100 hypothesized ambiguity vectors generated using the Fast Ambiguity Search Filter (FASF) routine [7]. The hypothesis conditional probability for each elemental filter was calculated in each of the GRMMAEs. A difference in probabilities was then calculated by subtracting the hypothesis conditional probabilities calculated in the traditional MMAE (GRMMAE with $\gamma = 1$) from those in the post-fit modified MMAE (GRMMAE with $\gamma = 0$). The probability difference between the two GRMMAEs was on the order of 10^{-12} and was likely due to machine precision,

rounding errors, and the slight differences in measurement matrices previously mentioned. Since the probability for the correct hypothesis is on the order of 1, a difference of 10^{-12} is essentially zero. Thus, this demonstrates that the GRMMAE with $\gamma = 1$ is equivalent to the GRMMAE with $\gamma = 0$ for this application, as predicted by the theory in Section 3.7.1.

4.3 Beta Dominance

This dissertation repeatedly states that the GRMMAE is susceptible to β-dominance, described in Section 2.4.1, when the transformation matrix $\mathbf{T}_j(t_i)$ varies among the elemental filters. Derivation of this fact in Section 3.6 is one of the contributions of this research. This section will demonstrate that fact by showing an example in which the GRMMAE with $\gamma = 1$ (i.e., a traditional MMAE) functions properly while a GRMMAE with $\gamma = 0$ (i.e., a post-fit modified MMAE) experiences β-dominance.

Recall from Section 3.6 that the multivariate Gaussian density function used to calculate the hypothesis conditional probabilities in the GRMMAE is given by

$$\frac{1}{(2\pi)^{\frac{m}{2}} \operatorname{abs}\left(|\mathbf{T}_j(t_i)|\right)|\mathbf{A}_j\left(t_i^-\right)|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}\mathbf{r}_j^T\left(t_i^-\right)\mathbf{A}_j^{-1}\left(t_i^-\right)\mathbf{r}_j\left(t_i^-\right)\right\} \tag{4.16}$$

This density is composed of a "beta term" and an "exponential term," respectively, given by

$$\beta_j^*(t_i) = \frac{1}{(2\pi)^{\frac{m}{2}} \operatorname{abs}\left(|\mathbf{T}_j(t_i)|\right)|\mathbf{A}_j\left(t_i^-\right)|^{\frac{1}{2}}}$$

and

$$\exp\left\{-\frac{1}{2}\mathbf{r}_j^T\left(t_i^-\right)\mathbf{A}_j^{-1}\left(t_i^-\right)\mathbf{r}_j\left(t_i^-\right)\right\} \tag{4.17}$$

In Equation (4.17), assuming linear system models are used, only the exponential term depends upon the measurements. The beta term is based solely on the system models. As such, it is completely pre-computable before any measurements are taken. Therefore, to ensure the GRMMAE converges to the correct elemental filter, the hypothesis conditional

probability calculation must be driven by the exponential term and not the beta term. In some applications, the system models are such that the beta terms have a wide separation between them, and the residuals are similar enough to cause the exponential terms to have similar magnitudes. When this happens, the beta terms dominate the hypothesis conditional probability calculation, and incorrect results may be achieved.

An example of β-dominance is created here to illustrate the GRMMAE's susceptibility to this phenomenon and to explain β-dominance further. The truth model of Equations (4.3) through (4.12) was used with a slight change to Equation (4.10). For this example, the true noise covariance matrices were

$$\mathbf{Q}_d(t_i) = 2.5 \times \begin{bmatrix} 0.0454 & 0.0838 & 0.0646 \\ 0.0838 & 0.1548 & 0.1193 \\ 0.0646 & 0.1193 & 0.0920 \end{bmatrix} \qquad \mathbf{R}(t_i) = \begin{bmatrix} 0.0025 & 0 \\ 0 & 0.0025 \end{bmatrix} \tag{4.18}$$

The measurement noise covariance in Equation (4.18) was the same as that in High $\mathbf{Q}_d$/Low $\mathbf{R}$ scenario of Equation (4.10), while the equivalent discrete-time noise covariance was 2.5 times that in the High $\mathbf{Q}_d$/Low $\mathbf{R}$ scenario of Equation (4.10). The change to the equivalent discrete-time noise covariance was made to facilitate creation of β-dominance.

The application used for this example is the unknown $\mathbf{Q}_d$ application, similar to Section 4.1.3. Three elemental filters were built based upon three hypotheses for the equivalent discrete-time noise

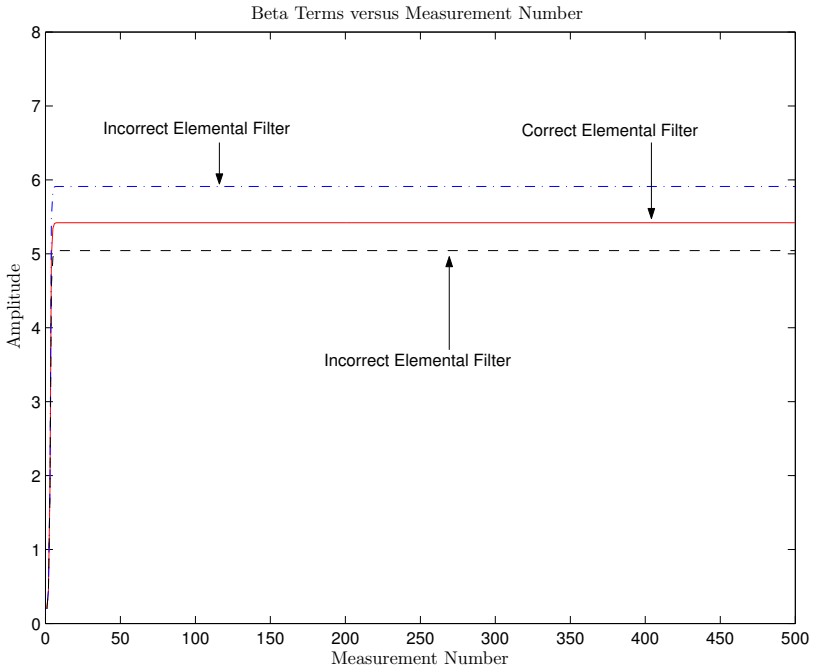
$$2\mathbf{Q}_d(t_i) \qquad 2.5\mathbf{Q}_d(t_i) \qquad 3\mathbf{Q}_d(t_i)$$

where $\mathbf{Q}_d$ is the matrix given in the High $\mathbf{Q}_d$/Low $\mathbf{R}$ scenario of Equation (4.10). Since there is only a small change between these three hypotheses, it is expected that the GRMMAE will converge to the correct solution relatively slowly. It is also expected that, if there is a large separation between the beta terms in the elemental filters, then the GRMMAE will experience β-dominance. Here, "a large separation" is most easily explained

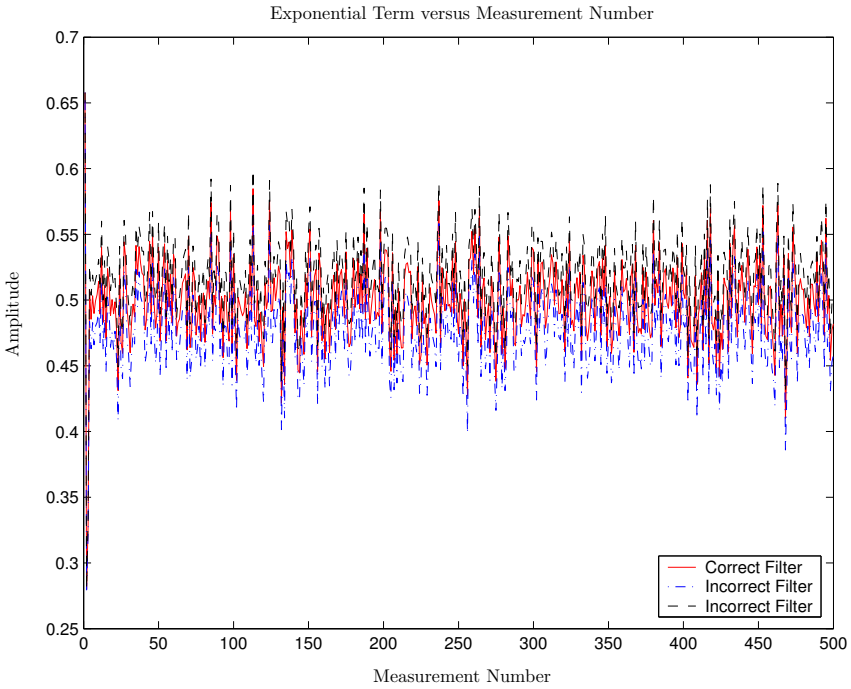
graphically. A GRMMAE was simulated 100 times with $\gamma = 1$ (i.e., a traditional MMAE), and each simulation consisted of 500 measurement epochs. Figure 4.12 is a plot of the sample averages of the exponential terms and the beta terms for these simulations, and Figure 4.13 gives an expanded plot of the exponential terms. Since the beta terms do not depend upon the measurements, they are the same in all 100 simulations.

In Figure 4.12 the exponential terms all have similar magnitudes. Recall from Section 2.4.3, that it is desirable for $\mathbf{r}_k^T\left(t_i^-\right)\mathbf{A}_k^{-1}\left(t_i^-\right)\mathbf{r}_k\left(t_i^-\right)$ to be on the order of the measurement dimension (2 in this example) [46]. Thus, the exponential term for the correct filter should be on the order of $\exp\left(-1\right) = 0.368$, which it is in this example. However, the exponential terms from the other elemental filters are on that same order. As a result, the GRMMAE with $\gamma = 1$ converges rather slowly to the correct hypothesis (as will be shown in Figure 4.17). Also note that the beta terms in Figure 4.12 are relatively small and closely spaced. Thus, it is expected that this GRMMAE (with $\gamma = 1$) will not be susceptible to β-dominance. To continue with this illustration, Figure 4.14 shows a plot of the product of the beta terms and the sample average exponential terms at each measurement epoch. In Figure 4.14, part (a) shows the full plot while part (b) expands on the early portion of the run to show more detail. From Figure 4.14, the products of the beta and sample average exponential terms have similar magnitudes for all three elemental filters. Therefore, the GRMMAE hypothesis conditional probability calculation is not dominated by the beta terms. Rather, the GRMMAE is able to use both residuals (i.e., the exponential terms) and the prior probabilities in the hypothesis conditional probability calculation.

In contrast, the same 100 measurement sets were used to repeat this example for a GRMMAE with $\gamma = 0$ (i.e., a post-fit modified MMAE). In addition to using the same measurements, the same GRMMAE was used with the exception that the γ design parameter was changed from 1 to 0. Thus, a direct comparison between GRMMAEs with $\gamma = 1$ and $\gamma = 0$ was made. Figure 4.15 shows a plot of the beta and the sample average exponential terms for the GRMMAE with $\gamma = 0$. Comparing Figures 4.12 and 4.15 shows that the sample average exponential terms were unchanged by the γ design parameter, as derived in Section 3.6. This confirms that the only dependence of the GRMMAE hypothesis conditional probability calculation on the γ design parameter is caused by the



(a) Beta Terms



(b) Exponential Terms

Figure 4.12 Beta and Sample Average Exponential Terms for GRMMAE with $\gamma = 1$ (Traditional MMAE)

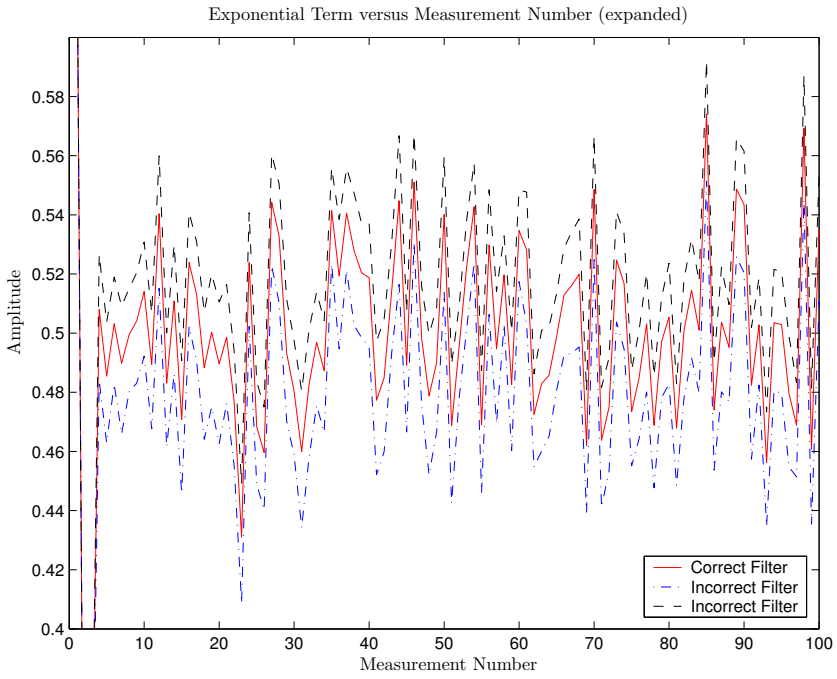
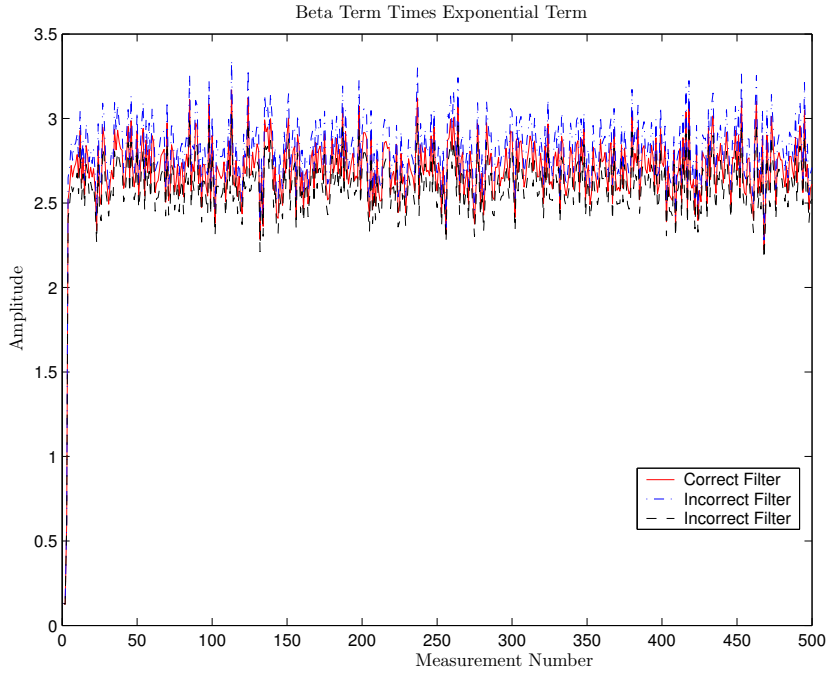


Figure 4.13 Sample Average Exponential Terms (expanded) for GRMMAE with $\gamma = 1$ (Traditional MMAE)

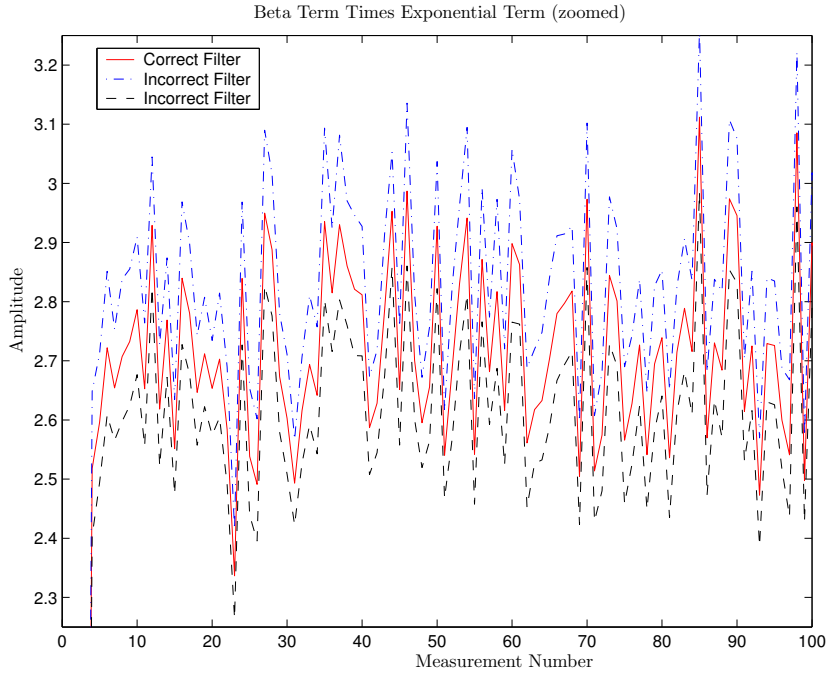
beta terms, demonstrating one of the key contributions of this dissertation. Notice that, in this example, the beta terms for the GRMMAE with $\gamma = 0$ are two orders of magnitude larger than the beta terms for the GRMMAE with $\gamma = 1$. This was caused by the determinant of the transformation matrix in the denominator of the beta term. When $\gamma = 1$, $|\mathbf{T}_j(t_i)| = 1 \ \forall j$ and it has no effect on the beta term. However, when $\gamma \neq 1$, the determinant has an effect on the beta terms. In this example, $\gamma = 0$ causes $\text{abs}\big(|\mathbf{T}_j(t_i)|\big)$ to increase greatly the magnitude of the beta terms[1].

In addition to increasing the magnitude of the beta terms, the separation between the beta terms was also increased when $\gamma = 0$. In this example, the increase in separation was enough to cause β-dominance. Figure 4.16, a plot of the product of the beta and sample average exponential terms, demonstrates this β-dominance. Unlike in Figure 4.14, the product of the beta and the sample average exponential terms shown in Figure 4.16 has sufficient separation between the hypotheses to cause the beta terms to dominate over

[1]It is not correct to say that $\text{abs}\big(|\mathbf{T}(t_i)|\big)$ *amplifies* the beta terms because the shape of the beta plots is changed in addition to the magnitude.

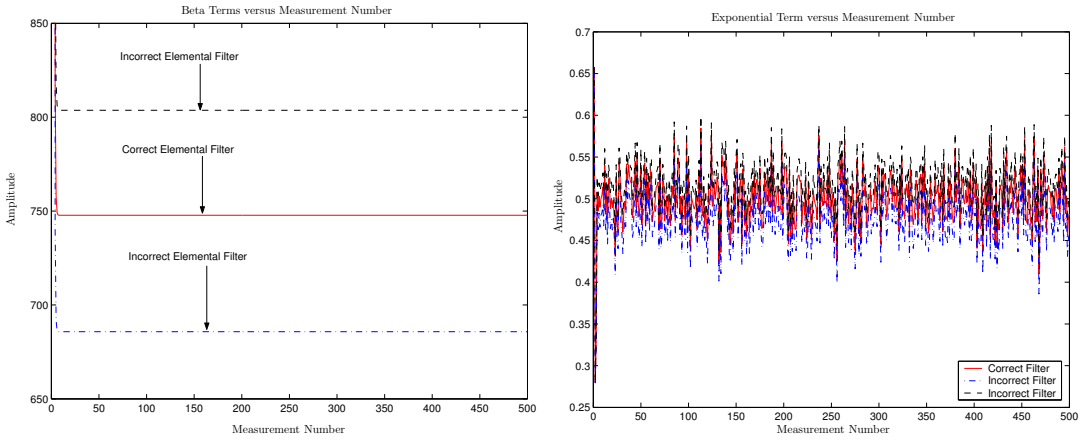


(a) Beta Term Times Exponential Term



(b) Beta Term Times Sample Average Exponential Term (expanded)

Figure 4.14 Product of Beta and Sample Average Exponential Terms for GRMMAE with $\gamma = 1$ (Traditional MMAE)



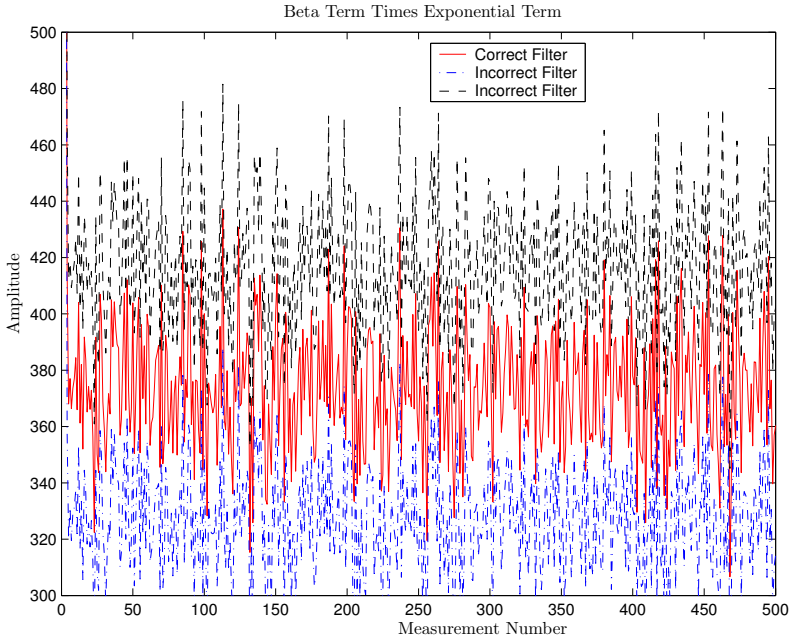
(a) Beta Terms

(b) Exponential Terms

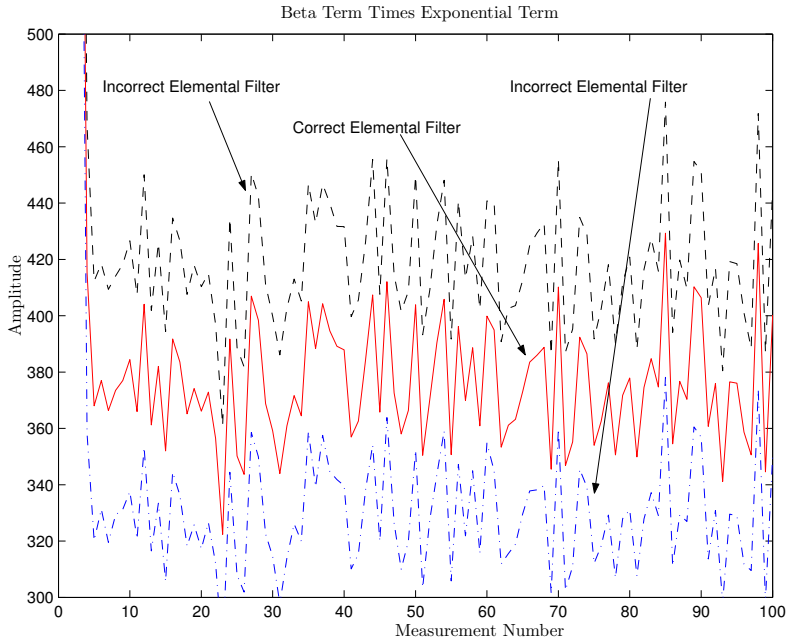
Figure 4.15 Beta and Sample Average Exponential Terms for GRMMAE with $\gamma = 0$ (Traditional MMAE)

the exponential terms. At every measurement epoch in Figure 4.16 (b), the product of the beta and sample average exponential terms for elemental filter 3 was larger than the same product for the other filters. This was caused by the dominant influence of the beta term for elemental filter 3. The root cause of this was the separation between the beta terms created by $\text{abs}\big(|\mathbf{T}_j(t_i)|\big)$.

As final evidence that the GRMMAE with $\gamma = 0$ suffered from β-dominance while the one with $\gamma = 1$ did not, the sample mean hypothesis conditional probabilities from the 100 Monte Carlo runs are shown in Figure 4.17. Part (a) of the figure shows the sample average hypothesis conditional probabilities for the GRMMAE with $\gamma = 1$ (i.e., a traditional MMAE). In this figure, the probability for the correct elemental filter slowly converged toward 1. It is reasonable to assume that this probability would have converged to 1 if the simulation had been allowed to continue longer. Part (b) of the figure, on the other hand, shows the hypothesis conditional probabilities for the GRMMAE with $\gamma = 0$ (i.e., a post-fit modified MMAE). In that figure, the probability for elemental filter 3, an incorrect hypothesis, very quickly converged to 1, while the probability for the correct hypothesis converged to 0. This occurred because the products of the beta and sample average exponential terms, shown in Figure 4.16, for the hypotheses were separated

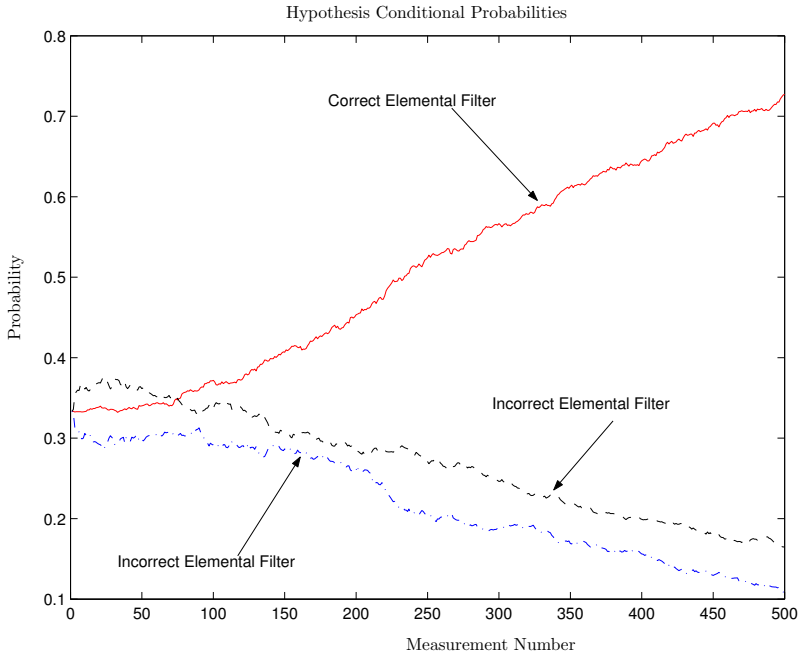


(a) Beta Term Times Exponential Term

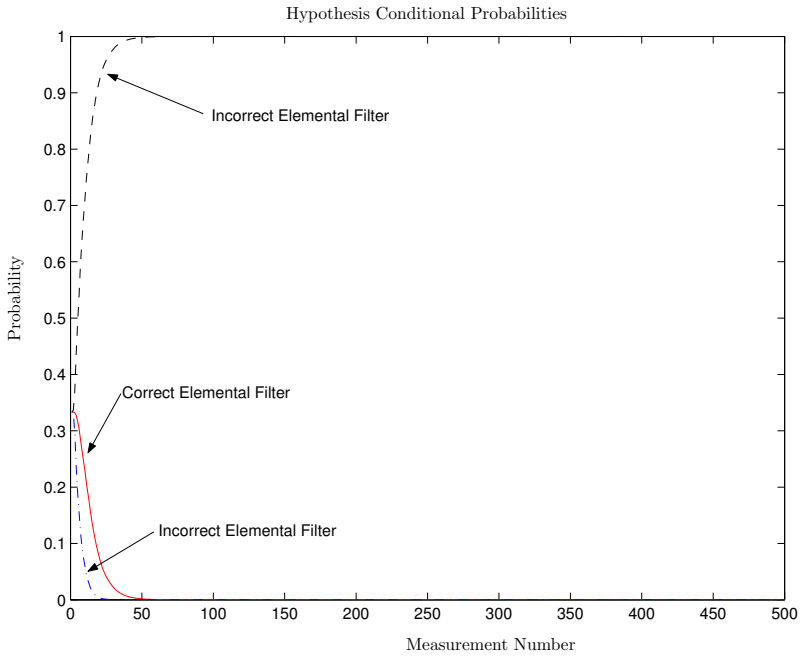


(b) Beta Term Times Exponential Term (expanded)

Figure 4.16 Product of Beta and Sample Average Exponential Terms for GRMMAE with $\gamma = 0$ (Post-Fit Modified MMAE)



(a) Probabilities for GRMMAE with $\gamma = 1$



(b) Probabilities for GRMMAE with $\gamma = 0$

Figure 4.17 Hypothesis Conditional Probability Comparison for GRMMAEs with $\gamma = 1$ and $\gamma = 0$

sufficiently to allow the beta terms to dominate the hypothesis conditional probability calculation. In this example, with $\gamma = 0$, elemental filter 3 had the largest beta term, as shown in Figure 4.12. Therefore, the β-dominance caused the GRMMAE with $\gamma = 0$ to converge to elemental filter 3, despite the fact that this was an incorrect hypothesis.

4.4 Effect of γ on Beta Terms

The previous section demonstrated that the design parameter γ changes the magnitude and the shape of the $\beta_j^*(t_i)$ plots. A mathematical derivation of the effects of γ on $\beta_j^*(t_i)$ does not support the primary goal of this research and, therefore, is not pursued. However, certain properties were observed and are demonstrated here. These observations are given without derivation or proof in support of a recommendation for future research given in Chapter V.

First, it was noted in the previous section that γ changes the shape of the $\beta_j^*(t_i)$ plots. This occurs only in the transient portion of the curve before the elemental filters reach steady state. For most γ values, $\beta_j^*(t_i)$ *increases* to a steady state value during the transient portion of the curve. This behavior is seen in Figure 4.12 (a). This shape makes sense because the residual covariance $\mathbf{A}_j\left(t_i^-\right)$, which appears in the denominator of $\beta_j^*(t_i)$, typically decreases as the elemental filter reaches steady state. However, at some point as γ approaches zero, the curve flips such that $\beta_j^*(t_i)$ *decreases* to a steady state value during the transient portion of the curve, as seen in Figure 4.15 (a). This flip appears to occur when $\text{abs}\left(|\mathbf{T}_j(t_i)|\right)$ begins to dominate over $|\mathbf{A}_j\left(t_i^-\right)|^{1/2}$ in the denominator of the beta terms. Analogous to β-dominance, this effect could be termed "**T**-dominance." This **T**-dominance effect appears to occur when γ is decreased below a threshold value which causes $\text{abs}\left(|\mathbf{T}_j(t_i)|\right)$ to become small enough to dominate over $|\mathbf{A}_j\left(t_i^-\right)|^{1/2}$ in the beta term. Confirmation of this observation and theoretical determination of the threshold γ are not pursued here, since they do not support the primary goal of this research. However, if a theoretical threshold for γ is determined, it may be possible to use the GRMMAE to correct for β-dominance rather than using the traditional method of β stripping. Alternately, this flip in the $\beta_j^*(t_i)$ curve may be related to the unexpected shape of Figure 4.2, in which the generalized residual covariance *increased* to a steady state value when $\gamma = 0$ (i.e.,

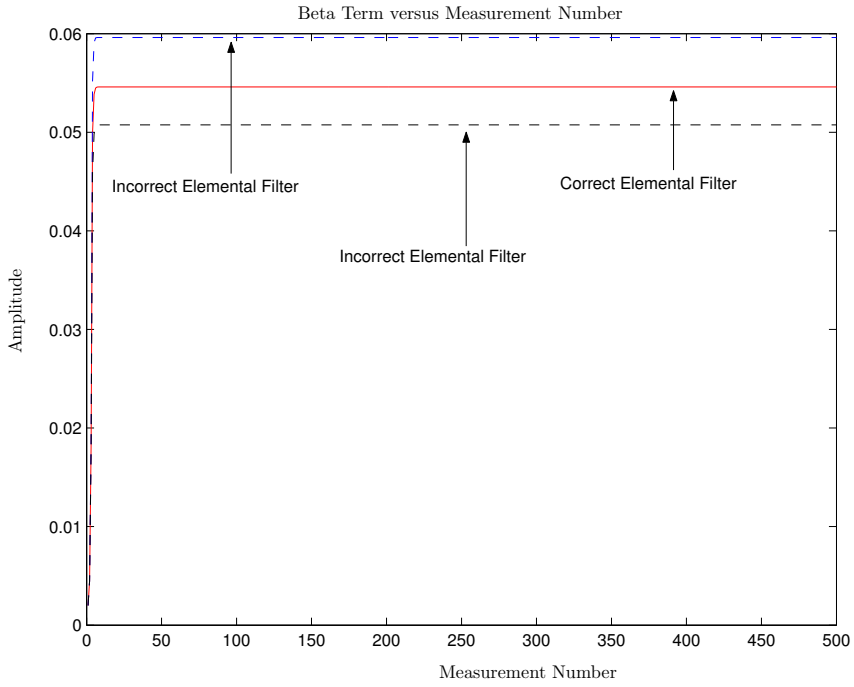
post-fit residual) was used. Again, there may be a threshold γ which caused that behavior. Finding this γ threshold would support development of a method to use γ to correct for β-dominance.

Secondly, it was observed that, as γ becomes large in absolute value, the beta terms generally become small. To demonstrate one example of this, the analysis from the previous section was re-run for γ equal 100, 10, -10, and -100, and the beta terms were plotted in Figure 4.18. Since the beta terms do not vary from one Monte Carlo run to the next, there was no need to run multiple simulations for this example. The first observation from Figure 4.18 is that the beta plots were so similar that they appear the same for numbers of equal magnitude but opposite signs (i.e., 100 and -100 produce nearly identical plots). The second and more important observation is that larger (magnitude) γ values led to smaller $\beta_j^*(t_i)$ values. This inverse relationship is also demonstrated in Figure 4.18.

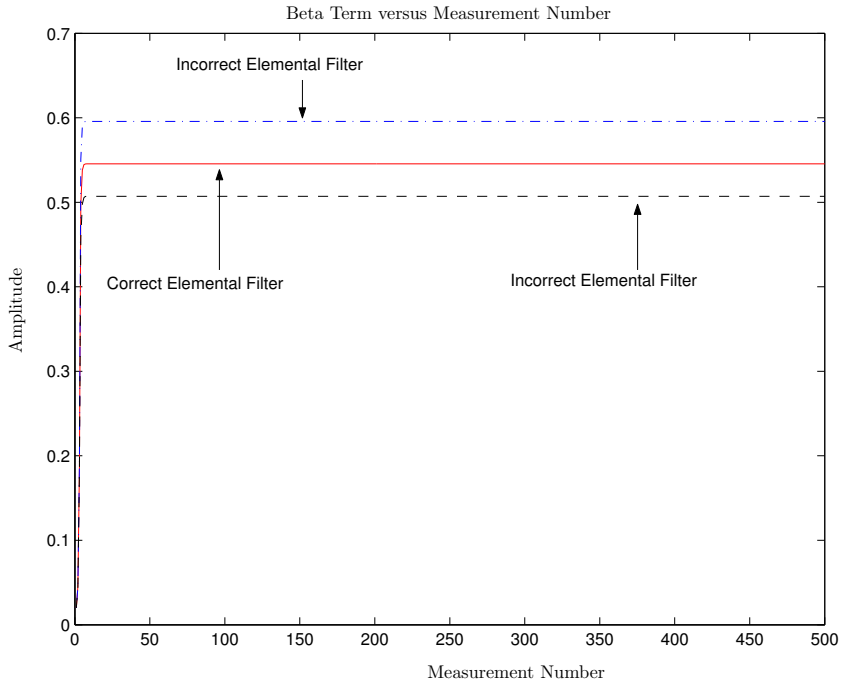
From Figure 4.18, it appears that increasing γ by an order of magnitude causes the beta terms to decrease by an order of magnitude. This observation is further supported by comparing Figures 4.18 and 4.12. However, Figure 4.15 clearly shows that relationship does not hold for all γ. From Figures 4.12 and 4.15, decreasing γ from 1 to 0 caused the beta terms to increase by 2 orders of magnitude. Thus, some type of singular behavior is observed as $\gamma \to 0$. Again, this observation does not support the primary goal of this research. Therefore, a mathematical description of this behavior is not pursued in this dissertation. However, a mathematical description of the relationship between γ and the beta terms is probably the key to using the GRMMAE, with an appropriate γ, to correct β-dominance, rather than using the ad hoc β stripping.

4.5 Scalar Penalty Increase

Section 2.4.5 mentioned scalar penalty increase as an alternative to the GRMMAE to cause faster MMAE convergence to the best parameter [31]. Although scalar penalty increase is not a research contribution of this dissertation, an example is presented here, since it is an alternative to the GRMMAE. Monte Carlo simulations of a system with an unknown measurement bias were used to demonstrate scalar penalty increase. With the



(a) GRMMAE with $\gamma = \pm 100$ ($+100$ and -100 lines overlap)



(b) GRMMAE with $\gamma = \pm 10$ ($+10$ and -10 lines overlap)

Figure 4.18 Beta Terms for GRMMAEs with $\gamma = \pm 100$ and $\gamma = \pm 10$

exception of the noise covariance matrices, the truth model from Section 4.1.1 was used for these simulations. The noise covariance matrices for these simulations were

$$\mathbf{Q}_d(t_i) = \begin{bmatrix} 0.0454 & 0.0838 & 0.0646 \\ 0.0838 & 0.1548 & 0.1193 \\ 0.0646 & 0.1193 & 0.0920 \end{bmatrix} \times 10^{-4} \qquad \mathbf{R}(t_i) = \begin{bmatrix} 1.5625 & 0 \\ 0 & 1.5625 \end{bmatrix} \tag{4.19}$$

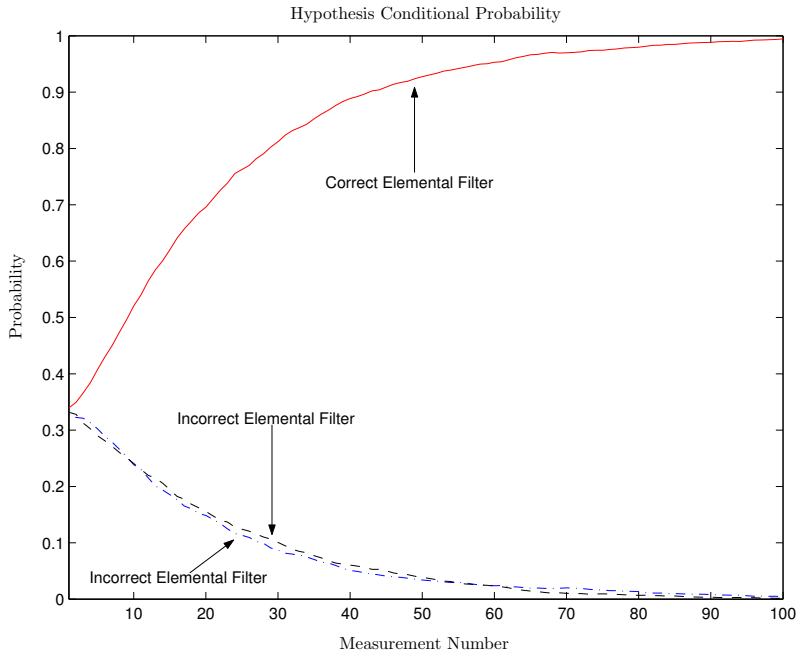
As before, three elemental filters were used. The three hypotheses for the measurement bias were

$$\mathbf{n}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \mathbf{n}_2 = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \qquad \mathbf{n}_3 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{4.20}$$

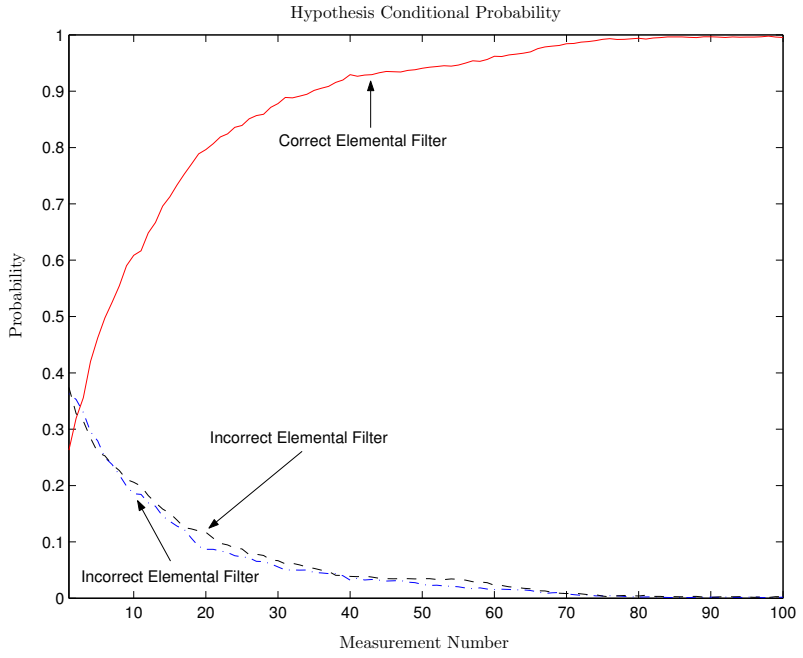
The true bias was $\mathbf{n}_1$. These hypotheses were changed from those in Section 4.1.2 to create a more challenging scenario for the MMAE. By making them very similar, the residuals from the three elemental filters will also be similar. This causes the MMAE to converge to the correct hypothesis more slowly, allowing a better opportunity to demonstrate the effect of scalar penalty increase.

The system was simulated 1000 times each for a traditional MMAE (i.e., GRMMAE with $\gamma = 1$) with and without scalar penalty increase. Figure 4.19 (a) shows the sample average hypothesis conditional probabilities calculated for the traditional MMAE without scalar penalty increase. Figure 4.19 (b) shows the hypothesis conditional probability for the traditional MMAE with scalar penalty increase. In the MMAE with scalar penalty increase, the exponential term in the Gaussian probability densities of Equation (2.41) was changed from $\exp\{ -\frac{1}{2}\mathbf{r}_k^T(t_i^-)\,\mathbf{A}_k^{-1}(t_i^-)\,\mathbf{r}_k(t_i^-)\}$ to $\exp\{ -2\mathbf{r}_k^T(t_i^-)\,\mathbf{A}_k^{-1}(t_i^-)\,\mathbf{r}_k(t_i^-)\}$.

Figure 4.19 demonstrates that the MMAE converged to the correct measurement bias vector regardless of whether scalar penalty increase was used or not. However, the MMAE with scalar penalty increase converged more quickly than the MMAE without scalar penalty increase. This simple example demonstrates that scalar penalty increase is useful to cause faster convergence, the goal the GRMMAE was originally proposed to



(a) Traditional MMAE Without Scalar Penalty Increase ($\alpha = 1/2$)



(b) Traditional MMAE With Scalar Penalty Increase ($\alpha = 2$)

Figure 4.19 Hypothesis Conditional Probability for a Traditional MMAE With and Without Scalar Penalty Increase

accomplish. The drawback to scalar penalty increase, however, is that it can lead to false alarms, as discussed in Section 2.4.5. Additionally, the "increased scalar" value must be chosen in an ad hoc manner. There is no method for choosing an optimal scalar value.

4.6 Chapter Summary

This chapter provided simulations to verify the theory derived in Chapter III. First, the generalized residual vector and covariance matrix equations from Section 3.2 were verified using a series of simple simulations. Next, since GPS carrier-phase integer ambiguity resolution is the application for which the GRMMAE was originally designed, double difference carrier-phase measurements were used to demonstrate that the GRMMAE, when applied to a measurement bias problem (same $\mathbf{T}(t_i)$ in all elemental filters), is independent of the γ design parameter. Then, the unknown $\mathbf{Q}_d$ application (differing $\mathbf{T}_j(t_i)$ in elemental filters) was used to demonstrate that the MMAE is susceptible to the β-dominance effect. This example demonstrated how the γ parameter created the β-dominance effect. Next, observations concerning the effect of γ on the $\beta_j^*(t_i)$ terms were described. These observations are likely to serve as the base point for recommended future research into use of γ to correct for β-dominance. Finally, scalar penalty increase was demonstrated as an alternative to the GRMMAE. With the exception of Sections 4.4 and 4.5, all sections in this chapter demonstrate scientific contributions of this dissertation.

V. Conclusions and Recommendations

This dissertation has developed the mathematical foundation for using generalized residuals to calculate hypothesis conditional probabilities in a Multiple Model Adaptive Estimator (MMAE). The developed estimator is termed a Generalized Residual Multiple Model Adaptive Estimator (GRMMAE). This GRMMAE was inspired by the use of post-fit residuals in a modified MMAE for GPS carrier-phase integer ambiguity resolution. Previous research demonstrated that the modified MMAE can successfully resolve GPS carrier-phase integer ambiguities [17, 18]. However, that research did not provide a mathematical derivation of the modified MMAE or the effects of the post-fit residuals on the MMAE's performance. Thus, in an effort to expand post-fit residual use in MMAE applications, this research provided a derivation of the GRMMAE. Through this derivation, the effect of post-fit residuals on the GRMMAE's performance has been determined. This final chapter provides a chapter-by-chapter summary of the dissertation. Section 5.1 lists the research contributions made in this dissertation. Then, the chapter concludes with a recommendation for future research.

Chapter I of this dissertation presented an introduction to the problem of estimating system states in the presence of an unknown system parameter. It discussed common problems in which system parameters must be estimated in addition to estimating the system state, and gave the motivation for using the MMAE for parameter and state estimation. Finally, it discussed the modified MMAE, which uses post-fit residuals, as the motivation for this research.

Chapter II presented the background theory necessary to derive the GRMMAE. It presented derivations of both the linear and nonlinear (extended) Kalman filters. It gave a derivation of the traditional MMAE and described several common performance enhancements to the MMAE. Since this dissertation was motivated by a GPS application, it also presented a brief introduction to the GPS. It discussed precise GPS and the need to resolve carrier-phase integer ambiguities in order to estimate precise relative positions with GPS. Chapter II concluded by describing the modified MMAE for GPS carrier-phase integer ambiguity resolution. The modified MMAE is the technique which motivated this dissertation.

Chapter III presented the research contributions of this dissertation. This chapter provided the mathematical derivation of the GRMMAE. This GRMMAE is a new Multiple Model Adaptive Estimator designed such that, by varying a single, scalar design parameter, the GRMMAE can be equivalent to a traditional MMAE, a modified MMAE, or any other MMAE based on a linear combination of traditional and post-fit Kalman filter residuals. Chapter III began with the definition of the new generalized residual. Next, it gave a derivation of the properties of the generalized residual vector from the elemental Kalman filter based on the correct system parameter. After that, the properties for the generalized residual vectors from incorrect elemental filters were derived for several common MMAE applications. Chapter III showed that the generalized Kalman filter residual is related to the traditional Kalman filter residual through a simple transformation matrix $\mathbf{T}(t_i)$. The effect of this transformation matrix was described in Chapter III, and it gave a derivation of the GRMMAE hypothesis conditional probability formula based on the generalized residual. The relationship between the traditional MMAE hypothesis conditional probability formula and the GRMMAE hypothesis conditional probability formula was presented. This relationship demonstrated that the GRMMAE design parameter γ effects only the $\beta_j^*(t_i)$ term in the GRMMAE hypothesis conditional probability calculation. This finding is the most significant finding in this dissertation, since it means that the GRMMAE is artificially susceptible to β-dominance. Chapter III concluded by further developing the generalized hypothesis conditional probability formula under two conditions: 1) the transformation matrix $\mathbf{T}(t_i)$ is the same in all elemental filters; 2) the transformation matrix $\mathbf{T}(t_i)$ is not the same in all elemental filters. It showed that, under condition 1), the design parameter γ is cancelled from the hypothesis conditional probability calculation. Thus, under condition 1), the GRMMAE produces results equivalent to a traditional MMAE regardless of the value of γ. In that case, the GRMMAE is not susceptible to β-dominance. Under condition 2), on the other hand, the effect of γ on the hypothesis conditional probability calculation does not cancel. In that case, the GRMMAE is susceptible to β-dominance.

Chapter IV presented simulation results to demonstrate the theory derived in Chapter III. It began with a series of simulations to verify the equations for calculating the generalized residual vector and its covariance matrix from the traditional residual vector.

Next, the GRMMAE was used to demonstrate that applications with the same $\mathbf{T}(t_i)$ in all elemental filters are independent of γ. Since the GRMMAE derivation was motivated by the modified MMAE for GPS carrier-phase integer ambiguity resolution, GPS carrier-phase measurements were used for this example. Next, the susceptibility of the GRMMAE to β-dominance was demonstrated. Then, the effect of γ on the $\beta_j^*(t_i)$ terms was discussed. The observations about this effect represent a starting point for the recommended future research presented later in this chapter.

5.1 Research Contributions

The development of the Generalized Residual Multiple Model Adaptive Estimator is the primary contribution of this dissertation. The heart of this development is the derivation of the hypothesis conditional probability calculation based on generalized residuals. This development provides a simple, general design in which a single, scalar design parameter γ is used to specify a traditional MMAE, a modified MMAE, or any other MMAE using a linear combination of traditional and post-fit residuals. Several supporting contributions are developed in this dissertation as part of the GRMMAE foundation. These supporting contributions are derived in Chapter III and verified in Chapter IV. The supporting contributions are:

- Derived generalized residual properties
- Derived and described transformation $\mathbf{T}(t_i)$ from traditional to generalized residual
- Derived hypothesis conditional probability formula using generalized residual
- Derived relationship between traditional and generalized hypothesis conditional probability formulas
- Derived and demonstrated equivalence between MMAE and GRMMAE when elemental filters have same $\mathbf{T}(t_i)$
- Derived formula demonstrating GRMMAE susceptibility to β-dominance
- Demonstrated how $\gamma \neq 1$ causes β-dominance

Each of these contributions are summarized in the sections that follow.

5.1.1 Derived Generalized Residual Properties. Generalized residual properties were derived in Sections 3.2 and 3.3. Section 3.2 demonstrated that the generalized residuals from the correct elemental filter are a linear function of traditional residuals from that same elemental filter. Since the traditional residuals are white, Gaussian, and zero-mean [29], the generalized residuals from the correct filter are also white, Gaussian, and zero-mean. Additionally, given that $\mathbf{A}_k\left(t_i^-\right)$ is the covariance of the traditional residuals from the correct elemental filter [29, 30], the covariance of the generalized residuals from the correct elemental filter is given in Equation (3.4), repeated here for convenience

$$\mathbf{A}_k^*\left(t_i\right) \;=\; \mathbf{T}_k\left(t_i\right)\mathbf{A}_k\left(t_i^-\right)\mathbf{T}_k^T\left(t_i\right) \tag{3.4}$$

where $\mathbf{T}_k\left(t_i\right) = \left[\mathbf{I} \;-\; \left(1 \;-\; \gamma\right)\mathbf{H}_k\left(t_i\right)\mathbf{K}_k\left(t_i\right)\right]$ is the linear relationship between the traditional and generalized residuals and γ is defined in Equation (3.1).

In Section 3.3, the residual vector mean and covariance for an elemental filter with an unknown measurement bias were shown to be

$$\mathbf{m}_{r_j^*}\left(t_i\right) \;= \mathbf{T}\left(t_i\right)\left[\mathbf{H}\left(t_i\right)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right) - \widehat{\mathbf{x}}_j\left(t_i^-\right)\right) + \boldsymbol{\Delta}\mathbf{n}_{kj}\right] \tag{3.11}$$

$$\mathbf{A}_j^*\left(t_i\right) \;= \mathbf{T}\left(t_i\right)\mathbf{A}_k\left(t_i^-\right)\mathbf{T}^T\left(t_i\right) \tag{3.13}$$

In this case, it was demonstrated that $\mathbf{H}\left(t_i\right)$ and $\mathbf{K}\left(t_i\right)$ were the same in all elemental filters. Thus, $\mathbf{T}\left(t_i\right)$ was the same in all elemental filters and the subscript identifying the elemental filter was removed. Equations (3.11) and (3.13) were shown to apply to any application in which the transformation matrix $\mathbf{T}\left(t_i\right)$ does not vary among the elemental filters. For an unknown $\mathbf{Q}_d\left(t_i\right)$, $\mathbf{F}\left(t\right)$ (and thus $\boldsymbol{\Phi}\left(t_{i+1}, t_i\right)$), $\mathbf{R}\left(t_i\right)$, or any other application for which the transformation matrix $\mathbf{T}\left(t_i\right)$ varies among the elemental filters, the generalized residual mean and covariance for an incorrect filter were shown to be

$$\mathbf{m}_{\mathbf{r}_j^*}\left(t_i\right) = \mathbf{T}_j\left(t_i\right)\mathbf{H}\left(t_i\right)\left[\widehat{\mathbf{x}}_k\left(t_i^-\right) \;-\; \widehat{\mathbf{x}}_j\left(t_i^-\right)\right] \tag{3.15}$$

$$\mathbf{A}_j^*\left(t_i\right) \;= \mathbf{T}_j\left(t_i\right)\mathbf{A}_k\left(t_i^-\right)\mathbf{T}_j^T\left(t_i\right)$$

These residual properties support the main contribution of providing a mathematical foundation for the use of generalized residuals in an MMAE. Similar derivations of traditional residual properties for a traditional MMAE are found in [16].

5.1.2 Derived and Described Transformation From Traditional to Generalized Residual $\mathbf{T}(t_i)$. The transformation from traditional to generalized residuals is given in Equation (3.2). It is

$$\mathbf{r}^*(t_i) \;=\; \mathbf{T}(t_i)\,\mathbf{r}\left(t_i^-\right) \tag{3.2}$$

Section 3.4 demonstrates that the effect of this transformation is to scale and rotate the residual vector in the measurement space. Additionally, it was stated that, in general, the matrix $\mathbf{T}(t_i)$ is not idempotent, so it is a transformation and not a projection. These contributions support the primary contribution by deriving and characterizing the effect of generalized residuals on the GRMMAE hypothesis conditional probability calculation.

5.1.3 Derived Hypothesis Conditional Probability Formula Using Generalized Residual. The derivation of the hypothesis conditional probability formula using generalized residuals is given in Section 3.5. This derivation is the heart of this dissertation. Previous research implemented the MMAE with post-fit residuals [17, 18]. However, that research did not mathematically derive the hypothesis conditional probability formula, nor did it mathematically characterize the effect of using non-traditional residuals in the MMAE. In Section 3.5, this dissertation provided a mathematical derivation of the GRMMAE hypothesis conditional probability calculation. It was shown that, by setting $\gamma = 0$, the GRMMAE is equivalent to the modified MMAE using post-fit residuals. However, the GRMMAE is more general than the modified MMAE. By setting $\gamma = 1$, the GRMMAE is equivalent to a traditional MMAE. Additionally, by setting γ equal to any other real, scalar number, the GRMMAE implements an infinite number of new MMAE's not encountered in scientific literature to date. Thus, the GRMMAE derivation provides a mathematical foundation for the modified (i.e., post-fit residual) MMAE which also applies to a more general class of MMAE's. This mathematical foundation is not found in scientific literature to date and is the primary contribution of this dissertation.

5.1.4 Derived Relationship Between Traditional and Generalized Hypothesis Conditional Probability Formulas. Using the GRMMAE mathematical foundation, the effect of generalized residuals in the MMAE is derived in Section 3.6. This derivation demonstrates that using $\gamma \neq 1$ affects only the $\beta_j^*(t_i)$ terms in the GRMMAE hypothesis conditional probability calculation. The specific choice of γ will increase or decrease the magnitudes of the $\beta_j^*(t_i)$ terms. Thus, γ is similar to a tuning parameter. It is used to "tune" the relative weight of the $\beta_j^*(t_i)$ terms compared to the $\exp\left\{-1/2\mathbf{r}_j^T\left(t_i^-\right)\mathbf{A}_j^{-1}\left(t_i^-\right)\mathbf{r}_j\left(t_i^-\right)\right\}$ terms in the hypothesis conditional probability calculation. However, unlike the exponential terms, the $\beta_j^*(t_i)$ terms are based solely on the elemental filter models and are unaffected by the real-world measurements (assuming the elemental filter models are linear). Therefore, using γ to artificially increase the $\beta_j^*(t_i)$ terms is ill-advised, as it may lead to β-dominance and incorrect results in the GRMMAE. This finding is a key contribution, as it points out a potentially fatal flaw if other researchers attempt to use post-fit residuals (or other generalized residuals) in the MMAE. Specifically, as emphasized throughout this dissertation, any generalized residual except the traditional residual (i.e., $\gamma = 1$) should not be used in the GRMMAE except in the rare case when the transformation matrix $\mathbf{T}_j(t_i)$ is the same in all elemental filters, or potentially to compensate for β-dominance encountered using a traditional MMAE.

5.1.5 Derived Equivalence Between MMAE and GRMMAE When Elemental Filter Have Same $\mathbf{T}(t_i)$. The previous subsection mentioned the case when all elemental filters have the same transformation matrix $\mathbf{T}(t_i)$. This special case applies to the GPS carrier-phase integer ambiguity resolution application as well as other applications in which an unknown measurement bias is estimated. Section 3.7.1 describes this special case, demonstrating that the hypothesis conditional probability for these applications is independent of the γ design parameter. Specifically, this explains why the modified (i.e., post-fit residual) MMAE successfully resolved GPS carrier-phase integer ambiguities, as described in [17, 18], without being susceptible to β-dominance. This contribution furthers the understanding of the effect of generalized residuals on the GRMMAE.

5.1.6 Derived Formula Demonstrating GRMMAE Susceptibility to β-dominance.

Using the GRMMAE mathematical foundation, the hypothesis conditional probability formula, when the transformation matrix $\mathbf{T}_j(t_i)$ varies among the elemental filters, was derived. This formula is given in Equation (3.40), repeated here for convenience

$$p_k^*(t_i) = \frac{\exp\left\{q_k^*(t_i)\right\}}{\exp\left\{q_k^*(t_i)\right\} + \sum\limits_{\substack{j=1 \\ j\neq k}}^{K} \text{abs}\left(\frac{|\mathbf{T}_k(t_i)|}{|\mathbf{T}_j(t_i)|}\right)\frac{p_j^*(t_{i-1})}{p_k^*(t_{i-1})}\exp\left\{q_j^*(t_i)\right\}} \tag{3.40}$$

where

$$\begin{aligned} q_j^*(t_i) = &-\frac{1}{2}\left[\mathbf{r}_k\left(t_i^-\right) + \mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right) - \widehat{\mathbf{x}}_j\left(t_i^-\right)\right)\right]^T \mathbf{A}_k^{-1}\left(t_i^-\right)\left[\mathbf{r}_k\left(t_i^-\right)\right. \\ &\left. + \mathbf{H}(t_i)\left(\widehat{\mathbf{x}}_k\left(t_i^-\right) - \widehat{\mathbf{x}}_j\left(t_i^-\right)\right)\right] \qquad \forall\ j \end{aligned}$$

Equation (3.40) implicitly shows that the γ design parameter, which is part of the transformation matrices $\mathbf{T}_k(t_i)$ and $\mathbf{T}_j(t_i)$, affects the $\beta_j^*(t_i)$ terms, but not the exponential terms. This is a direct cause of the GRMMAE's increased vulnerability to β-dominance if $\gamma = 1$ (i.e., the traditional residual vector) is not used. As described in Section 5.1.4, this conclusion is an important contribution of this dissertation.

5.1.7 Demonstrated How $\gamma \neq 1$ Causes β-dominance. Section 4.3 demonstrated in detail how using $\gamma \neq 1$ causes the GRMMAE to be susceptible to β-dominance. It demonstrated that an improper choice for the γ parameter causes the $\beta_j^*(t_i)$ terms to become large and increases the Euclidean distance between $\beta_j^*(t_i)$ and $\beta_k^*(t_i)$ for $j \neq k$. The conclusion from this section is that, in general, $\gamma \neq 1$ should not be used in the GRMMAE as it is currently implemented. The exception to this conclusion occurs when the transformation matrix $\mathbf{T}(t_i)$ does not vary among the elemental filter. This is a key finding of this dissertation, as it demonstrates that the practice of using post-fit residuals as in [17, 18] is not generally sound.

5.2 Recommendation for Future Research

Due to the GRMMAE's susceptibility to β-dominance, the GRMMAE is not recommended for widespread application. However, there are certain applications in which β-dominance occurs with a traditional MMAE. Most notably, β-dominance has been observed and documented in the failure detection application [32, 45]. The standard method to address this problem is to use β stripping as described in Section 2.4.1.

While β stripping is an effective method to overcome β-dominance, it is ad hoc. There is no claim or guarantee that β stripping is optimal. However, the GRMMAE provides a mathematical formulation in which an appropriate optimality criterion can be defined, and a γ value chosen to achieve optimal performance. Therefore, a research effort into using the GRMMAE in place of β stripping to address β-dominance is recommended as a follow-on to this dissertation. This follow-on research would use the design parameter γ to tune the $\beta_j^*(t_i)$ terms to be in consonance with the exponential terms. Such a study is envisioned as follows:

- Mathematically define the precise meaning for the $\beta_j^*(t_i)$ terms "to be in consonance with" the exponential terms. This definition will likely quantify the relationship between the magnitudes of the scalar $\beta_j^*(t_i)$ and exponential terms in the hypothesis conditional probability calculation. It will then probably set a threshold on the relative magnitudes such that the GRMMAE hypothesis conditional probabilities are not dominated by the $\beta_j^*(t_i)$ terms.
- Define an optimality condition, with γ as the optimization variable, to drive the $\beta_j^*(t_i)$ terms into consonance with the exponential terms.
- Derive an expression to solve for the optimal γ value. The optimal γ could be calculated at each measurement update, or one steady state value could be determined.
- Finally, the performance of the GRMMAE for β-dominance compensation should be compared with the performance of the traditional MMAE implementing β stripping to determine the performance gain attained using the optimal method.

Since the γ design parameter affects only the $\beta_j^*(t_i)$ term, there is potential it can be used to compensate for β-dominance. However, it is not clear that the use of γ will

enable full "controllability" of the β-terms on all elementals filters. Thus, implementing the suggested procedure is not a trivial effort. However, the suggested procedure is likely at least to provide an alternative to β stripping. Additionally, there is potential that it could provide improved performance and replace an ad hoc technique with an optimal one.

Bibliography

1. Apostol, T. *Calculus*, *I*. New York, London: Blaisdell Publishing Company, 1961.
2. Athans, M. and C. Chang. *Adapative Estimation and Parameter Identification Using Multiple Model Estimation Algorithm*. Technical Report ESD-TR-76-184, Defense Technical Information Center, Ft. Belvoir, VA: Massachusetts Institute of Technology, 1976.
3. Athans, M. et al. "The Stochastic Control of the F-8C Aircraft Using a Multiple Model Adaptive Control (MMAC) Method – Part I.: Equilibrium Flight," *IEEE Transactions on Automatic Control*, *AC-22*(5) (1977).
4. Bar-Shalom, Y. and X.R. Li. *Estimation and Tracking: Principles, Techniques, and Software*. Norwood, MA: Artech House, 1993.
5. Baram, Y. and Jr. N.R. Sandell. "Consistent Estimation of Finite Parameter Sets with Application to Linear Systems Identification," *IEEE Transactions on Automatic Control*, *AC-23*(3) (1978).
6. Baram, Y. and Jr. N.R. Sandell. "An Information Theoretic Approach to Dynamic System Modelling and Identification," *IEEE Transactions on Automatic Control*, *AC-23*(1) (1978).
7. Chen, D. and G. Lachapelle. "A Comparison of the FASF and Least-Squares Search Algorithms for Ambiguity Resolution on the Fly." *Proceedings of International Symposium on Kinematic Systems - KIS 94*. 241–253. August 30 - September 2 1994.
8. Eide, P. *Implementation and Demonstration of a Multiple Model Adaptive Estimator Failure Detection System for the F-16*. MS thesis, Air Force Institute of Technology, Wright-Patterson AFB, OH, December 1994.
9. Eide, P. and P. Maybeck. "An MMAE Failure Detection System for the F-16," *IEEE Transactions on Aerospace and Electronic Systems*, *32*(3):1125–1148 (July 1996).
10. Fisher, K. *Multiple Model Adaptive Estimation using Filter Spawning*. MS thesis, Air Force Institute of Technology, Wright-Patterson AFB, OH, March 1999.
11. Fisher, K. "Multiple Model Adaptive Estimation with Filter Spawning," *IEEE Transactions on Aerospace and Electronic Systems*, *38*(3) (July 2002).
12. Griffin, G. and P. Maybeck. "MMAE/MMAC Control for Bending with Multiple Uncertain Parameters," *IEEE Transactions on Aerospace and Electronic Systems*, *33*(3) (July 1997).
13. Gustafson, J. *Control of a Large Flexible Space Structure Using Multiple Model Adaptive Algorithms*. MS thesis, Air Force Institute of Technology, Wright-Patterson AFB, OH, December 1991.
14. Gustafson, J. and P. Maybeck. "Flexible Spacestructure Control Via Moving-Bank Multiple Model Algorithms," *IEEE Transactions on Aerospace and Electronic Systems*, *30*(3):750–757 (July 1994).

15. Hanlon, P. *Failure Identification Using Multiple Model Adaptive Estimation for the LAMBDA Flight Vehicle*. MS thesis, Air Force Institute of Technology, Air Force Institute of Technology, September 1992.

16. Hanlon, P. *Practical Implementation of Multiple Model Adaptive Estimation Using Neyman-Pearson Based Hypothesis Testing and Spectral Estimation Tools*. PhD dissertation, Air Force Institute of Technology, Wright-Patterson AFB, OH, September 1996.

17. Henderson, P., J. Raquet and P. Maybeck. "A Multiple Filter Approach for Precise Kinematic DGPS Positioning and Carrier-Phase Ambiguity Resolution," *NAVIGATION Journal of the Institute of Navigation*, *49*(3):149–160 (Fall 2002).

18. Henderson, P. *Development and Testing of a Multiple Filter Approach for Precise DGPS Positioning and Carrier-Phase Ambiguity Resolution*. MS thesis, Air Force Institute of Technology, Wright-Patterson AFB,OH, March 2001.

19. Hwang, P. *Application of GPS to Geodesy: A Combination Problem in Estimation and Large-Scale Hypothesis Testing*. MS thesis, Iowa State University, Ames, Iowa, 1983.

20. Hwang, P. "Kinematic GPS for Differential Positioning: Resolving Integer Ambiguities on the Fly," *Global Positioning System Papers Published in NAVIGATION*, *IV*:205–219 (1993).

21. Hwang, P. and R. Brown. "A Kalman Filter Approach to Precision GPS Geodesy," *Global Positioning System Papers Published in NAVIGATION*, *IV*:155–166 (1993).

22. Laxton, M. *Analysis and Simulation of a New Code Tracking Loop for GPS Multipath Mitigation*. MS thesis, Air Force Institute of Technology, Wright-Patterson AFB, OH, December 1996.

23. Li, X.R. and Y. Bar-Shalom. "Multiple Model Estimation with Variable Structure," *IEEE Transactions on Automatic Control*, *41*(4) (April 1996).

24. Li, X.R. et al. "Multiple-model Estimation with Variable Structure: Model-group Switching Algorithm." *Proceedings of the 36th IEEE Conference on Decision and Control*. December 1997.

25. Lund, E. *On-Line Discrimination and Estimation in Multiple Regime Systems*. PhD dissertation, Division of Engineering Cybernetics, Norwegian Institute of Technology, University of Trondheim, 1992.

26. Magill, D. "Optimal Adaptive Estimation of Sampled Stochastic Processes," *IEEE Transactions on Automatic Control*, *AC-10*:434–439 (1965).

27. Maybeck, P. *Adaptive Tracking of Maneuvering Targets Based on IR Image Data*. AGARD Lecture Series No. 166, July 1989.

28. Maybeck, P. *Moving-Bank Multiple Model Adaptive Estimation and Control Algorithms: An Evaluation*, *31*. New York: Academic Press, 1989.

29. Maybeck, P. *Stochastic Models, Estimation, and Control*, *1*. Navtech Book & Software Store, 1994.

30. Maybeck, P. *Stochastic Models, Estimation, and Control*, *2*. Navtech Book & Software Store, 1994.

31. Maybeck, P. and P. Hanlon. "Performance Enhancement of a Multiple Model Adaptive Estimator," *IEEE Transactions on Aerospace and Electronic Systems*, *31*(4):1240–1254 (October 1995).

32. Maybeck, P. and R. Stevens. "Reconfigurable Flight Control via Multiple Model Adaptive Control Methods," *IEEE Transactions on Aerospace and Electronic Systems*, *27*(3):470–480 (May 1991).

33. Menke, T. *Multiple Model Adaptive Estimation Applied to the VISTA F-16 with Actuator and Sensor Failures*. MS thesis, Air Force Institute of Technology, Wright-Patterson AFB, OH, June 1992.

34. Menke, T. and P. Maybeck. "Sensor/Actuator Failure Detection in the VISTA F-16 by Multiple Model Adaptive Estimation," *IEEE Transactions on Aerospace and Electronic Systems*, *AES-31*(4):1218–1228 (October 1995).

35. Misra, P. and P. Enge. *Global Positioning System Signals, Measurements, and Performance*, chapter 6, 209–257. Ganga-Jamuna Press, 2001.

36. Parkinson, B. and P. Enge. *Global Positioning System: Theory and Applications*, *164*. Progress in Astronautics and Aeronautics, chapter 1. Washington, DC: American Institute of Aeronautics and Astronautics, Inc., 1996.

37. Raquet, J. *Development of a Method for Kinematic GPS Carrier-Phase Ambiguity Resolution Using Multiple Reference Receivers*. PhD dissertation, The University of Calgary, May 1998.

38. Raquet, J., "EENG 533: Navigation Using The GPS Course Handouts." Air Force Institute of Technology, Spring 2001.

39. Raquet, J., "EENG 633: Advanced GPS Theory and Applications Class Notes." Air Force Institute of Technology, Summer 2001.

40. Reid, G. *Linear System Fundamentals Continuous and Discrete, Classic and Modern*. McGraw-Hill, 1983.

41. Schiller, G. *Control of a Large Space Structure Using Multiple Model Adaptive Estimation and Control Techniques*. MS thesis, Air Force Institute of Technology, Wright-Patterson AFB, OH, December 1993.

42. Schiller, G. and P. Maybeck. "Control of a Large Space Structure Using MMAE/MMAC Techniques," *IEEE Transactions on Aerospace and Electronic Systems*, *33*(4) (October 1997).

43. Stepaniak, M. *Multiple Model Adaptive Control of the VISTA F-16*. MS thesis, Air Force Institute of Technology, Wright-Patterson AFB, OH, December 1995.

44. Stepaniak, M. "MMAE-Based Control Redistribution Applied to the VISTA F-16," *IEEE Transactions on Aerospace and Electronic Systems*, *AES-34*(4) (October 1998).

45. Stevens, R. *Characterization of a Reconfigurable Multiple Model Adaptive Controller Using a STOL F-15 Model*. MS thesis, Air Force Institute of Technology, Wright-Patterson AFB, OH, December 1989.

46. Vasquez, J. *New Algorithms for Moving-Bank Multiple Model Adaptive Estimation*. PhD dissertation, Air Force Institute of Technology, 1998.

47. Young, R. and G. McGraw. "Fault Detection and Exclusion Using Normalized Solution Separation Methods." *Proceedings of the 15th International Technical Meeting of the Satellite Division of the Institute of Navigation: ION GPS 2002*. 566–578. September 2002.

Vita

Major Charles D. Ormsby earned a bachelor's degree in Electrical Engineering from Rose-Hulman Institute of Technology in 1992. He was commissioned through the ROTC program in May 1992 and entered active duty in February 1993. Major Ormsby's first assignment was to the National Air Intelligence Center (NAIC) at Wright Patterson AFB. There he worked as a microelectronics materials analyst and a foreign materiel exploitation project engineer. Following his assignment to NAIC, Major Ormsby attended the Air Force Institute of Technology (AFIT). There he earned a master's degree in Electrical Engineering with an emphasis on communications. From AFIT, Major Ormsby went to the 746^{th} Test Squadron at Holloman AFB. While at Holloman, he ran the Navigation Test and Evaluation Laboratory, providing hardware-in-the-loop simulation testing of military navigation systems. Additionally, he participated in a number of field tests of navigation systems in a GPS jamming environment. In 2000, Major Ormsby left Holloman to return to AFIT for the PhD program. During this second assignment to AFIT, Major Ormsby's major area of study has been Stochastic Estimation, Control, and Navigation. His minor area of study was Signal Processing. Following completion of the PhD program, Major Ormsby will be assigned to the Air Force Research Laboratory.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of the collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to an penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From – To)*
10-02-2003	**Doctoral Dissertation**	Jul 2002 – Oct 2003

4. TITLE AND SUBTITLE

GENERALIZED RESIDUAL MULTIPLE MODEL ADAPTIVE ESTIMATION OF PARAMETERS AND STATES

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Ormsby, Charles, D., Major, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)
Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way, Building 640
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT/DS/ENG/03-08

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Eric Lekberg
Joint Precision Approach and Landing System Program Manager
ESC/GAL
75 Vandenberg Drive
Hanscom AFB MA 01731-2103

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This dissertation develops a modification to the standard Multiple Model Adaptive Estimator (MMAE) which allows the use of a new "generalized residual" in the hypothesis conditional probability calculation. The generalized residual is a linear combination of traditional Kalman filter residuals and "post-fit" Kalman filter residuals which are calculated after measurement incorporation. This modified MMAE is termed a Generalized Residual Multiple Model Adaptive Estimator (GRMMAE). The dissertation provides a derivation of the hypothesis conditional probability formula which the GRMMAE uses to calculate probabilities that each elemental filter in the GRMMAE contains the correct parameter value. Through appropriate choice of a single scalar GRMMAE design parameter, the GRMMAE can be designed to be equivalent to a traditional MMAE, a post-fit residual modified MMAE, or any number of yet unused MMAEs. The original GRMMAE design goal was to choose the GRMMAE design parameter which caused the fastest GRMMAE convergence to the correct hypothesis. However, this dissertation demonstrates that the GRMMAE design parameter can lead to beta-dominance, a negative performance effect in the GRMMAE. This is a key contribution of this research as other researchers have previously suggested that the use of post-fit residuals may be advantageous in certain MMAE applications. This dissertation directly addresses the use of post-fit residuals by those researchers and demonstrates that, for their application, equivalent performance is achieved using a traditional MMAE.

15. SUBJECT TERMS

Kalman filtering, Adaptive systems, Adaptive filters, Multiple model adaptive estimation, MMAE, Global positioning system, GPS, Failure identification, Target tracking

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			John F. Raquet (ENG)
U	U	U	UU	147	19b. TELEPHONE NUMBER *(Include area code)* (937) 255-3636, ext 4580; e-mail: john.raquet@afit.edu